TIME SERIES DATA ANALYSIS AND SYNTHESIS FOR RESEARCH WATERSHEDS

112
ARS-S-~~76~~
January 1976

AGRICULTURAL RESEARCH SERVICE • U.S. DEPARTMENT OF AGRICULTURE

CONTENTS

Illustrations

Tables

TIME SERIES DATA ANALYSIS AND SYNTHESIS FOR RESEARCH WATERSHEDS

By W. M. Snyder[1]

ABSTRACT

Mathematical models of watershed flow processes can be used to simulate future runoff events. Such synthesized series of events form the basis for risk analysis in planning and design. Operation of the models requires generation of synthetic rainfall and other model inputs as stochastic series. This paper presents model forms for analysis and synthesis of information usually available from research watersheds.

INTRODUCTION

Fiering and Jackson[2] have presented the practical approach to generation of synthetic streamflows in an excellent and understandable tutorial monograph. A series of five computer programs has been designed to implement such generating schemes. Four of the programs are designed to examine historical data as an aid in choosing from among simple linear generating models. The fifth synthesizes streamflows for a particular model.

The programs are presented in terms of case studies on two watersheds operated by the Agricultural Research Service—Watkinsville, Ga., W-1, 19.2 acres; and Vero Beach (Taylor Creek), Fla., W-2, 98.6 square miles. Data for these watersheds are published in Hobbs[3] and subsequent reports. Published data for 1940–64, were available for the Watkinsville watershed. Data for 1956–66 were available for Taylor Creek watershed. In addition, mean monthly air temperatures at Lake Okeechobee, Fla., and Athens, Ga., were taken from Climatic Data publications of the U.S. Weather Bureau (now Weather Service). No further descriptions of the watersheds or the data will be given, since this report deals with development of numerical methodologies, not with analysis of any particular data set.

Data from the two Agricultural Research Service watersheds were used so that implementation of the methods and models suggested by Fiering and Jackson could be tested on watersheds smaller than those normally considered for water resource development. Such testing is a necessary part of the development of procedures to generate synthetic inputs for conceptual watershed models. Current emphasis in agricultural watershed modeling is on prediction of water runoff as the carrier of sediment and other materials. Ranges and extremes of water, sediment, and chemical runoff must be estimated by techniques of synthetic data generation.

This report presents some results of application of numerical methods not covered by Fiering and Jackson. Regression coefficients are estimated by both conventional and nonlinear least squares. Statistical frequency distributions with mathematically continuous, seasonally variable parameters are also estimated by nonlinear least squares.[4] Seasonally continuous regression coefficients and dis-

[1]Hydraulic engineer, Southeast Watershed Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Athens, Ga. 30601.

[2]Fiering, M. B., and Jackson, B. B. 1971. Synthetic streamflows. Am. Geophys. Union Water Resour. Monogr. 1, 98 pp.

[3]Hobbs, H. W. 1963. Hydrologic data for experimental agricultural watersheds in the United States. U.S. Dep. Agric. Misc. Publ. No. 945, pp. 8.2-1–8.2-4, 10.1-1–10.1-8.

[4]Snyder, W. M. 1972. Fitting of distribution functions by nonlinear least squares. Water Resour. Res. 8(6): 1423–1432, Snyder, W. M., and Wallace, J. R. 1974. Estimating the parameters of the log-normal distribution. Nordic Hydrol. 5(3):129–145.

tribution parameters are specified by a cyclic modification of the method of continuous parabolic interpolation.[5]

In the following presentation, no particular attempt is made to interpret the hydrologic implications of all outputs from the computer programs, but rather to illustrate how the programs provide a basis for decisionmaking on model structure, as discussed by Fiering and Jackson. The programs are not intended solely for use in their listed form. They do form core procedures which can be easily modified to incorporate other operations on data and other outputs as required by individual researchers. For example, many models will be based on other than monthly data, and modifications can be made to the programs to change the time base of the data.

For brevity of presentation, it is assumed that the reader is familiar with time-series analysis and synthesis from tutorial and applications perspectives such as presented by Fiering and Jackson.

AUTOCORRELATION AND CROSS CORRELATION

Consider the two series of variables: Y_i, $i=1, N$; and X_i, $i=1, N$. Here, Y_i through Y_N will represent monthly totals (or averages) of some quantity, such as watershed runoff, in chronological order from month 1 to month N of a record. Similarly, X_i's are such totals or averages. The X_i's represent independent, or causal, properties. The Y_i's represent dependent properties.

The end product of a data generation scheme is the estimation of one or more possible sets of the series Y_i. For example, one 50-year series of monthly runoff values might be synthesized, or 10 such series might be synthesized. The synthetic series are estimates of what might happen in the future. Assurance is needed that these predicted series are likely, to some degree of likeliness, usually based on one observed series of Y_i.

The likeliness of possible future series of values of Y_i is accomplished by extracting certain elements of information from the historical sequence and making certain that this same information is reproduced in the synthetic series. For example, the means and standard deviations of many synthetic series should have mean values very nearly the same as the historical series.

[5]Snyder, W. M. 1961. Continuous parabolic interpolation. J. Hydraul. Div., Proc. Am. Soc. Civ. Eng. 87(HY4): 99–111.

The starting point for a data-generating scheme must be a method of processing data so that information within the data can be recognized and quantified. It must be determined whether the Y_i's are serially related. Is there, for example, a relationship between $Y_{i=p}$ and $Y_{i=p+l}$, where l is some positive number of lags? Does this relation vary, and in particular does it tend to have a cyclic pattern as l increases from 1 to some number less than N? Such cyclic within-series patterns of information are usually determined by autocorrelation, which simply means the calculation of the correlation coefficient between $Y_{i=p}$ and $Y_{i=p+l}$ for various values of l.

When a supplemental data series, X_i, is available, relationships between this series and the series to be predicted should also be determined. If monthly rainfall is available, synchronous with the monthly runoff values, then it must be determined whether information in the runoff series can be cast into a rainfall-runoff relationship. Also, it must be determined whether this relationship is more important than the serial structure of the Y_i series in the prediction of Y_i.

Program I, listed in the appendix, is a simple program designed to generate the autocorrelation coefficients for one independent and one dependent variable and for the cross-correlation coefficients between the two. The dependent variable lags behind the independent variable, since it is obvious, for example, that runoff cannot anticipate future rainfall.

Figure 1 is a plot of output from program I for the Taylor Creek watershed. In the upper part of the figure, the auto- and cross-correlations for monthly rainfall and runoff are shown. In the lower part, the autocorrelation coefficient for monthly mean temperature and cross-correlation coefficients between temperature and rainfall and runoff are shown. Temperature is considered the independent variable. The coefficients were computed for all lags from 0 to 39.

The strong seasonal pattern of rainfall, runoff, and temperature is shown by the sinusoidal waves with a wave length of 12 lags in both the upper and lower portions of figure 1. Temperature shows the most distinct pattern, with correlations oscillating between values of −0.9 and +0.9. Rainfall also shows a distinct annual pattern, but values oscillate between about −0.4 and +0.4, showing a much greater year-to-year variability than temperature. Monthly runoff has the lowest autocorrelation coefficients, and hence the greatest year-to-year variability.

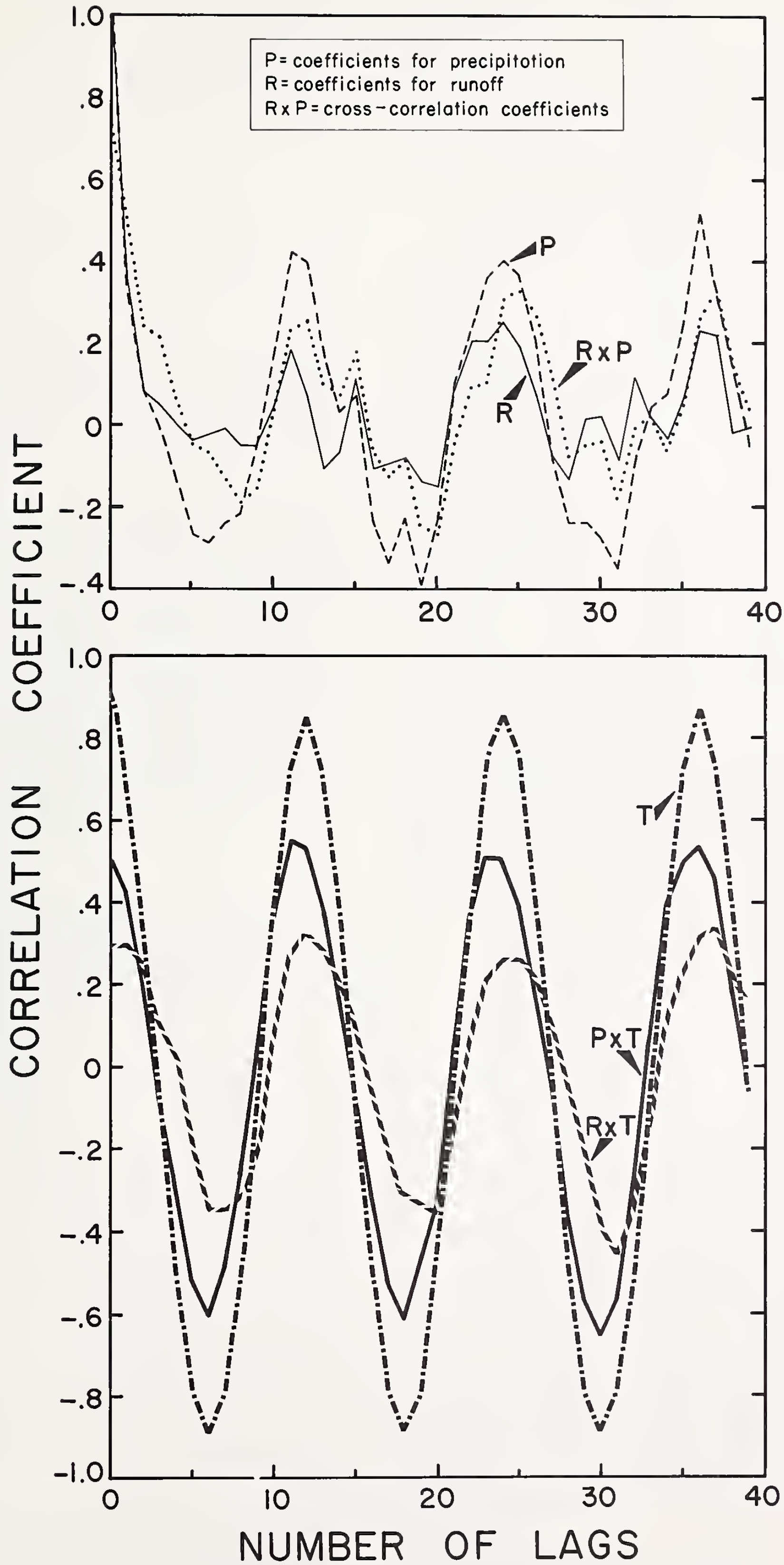

FIGURE 1.—Auto- and cross-correlation coefficients for Taylor Creek watershed.

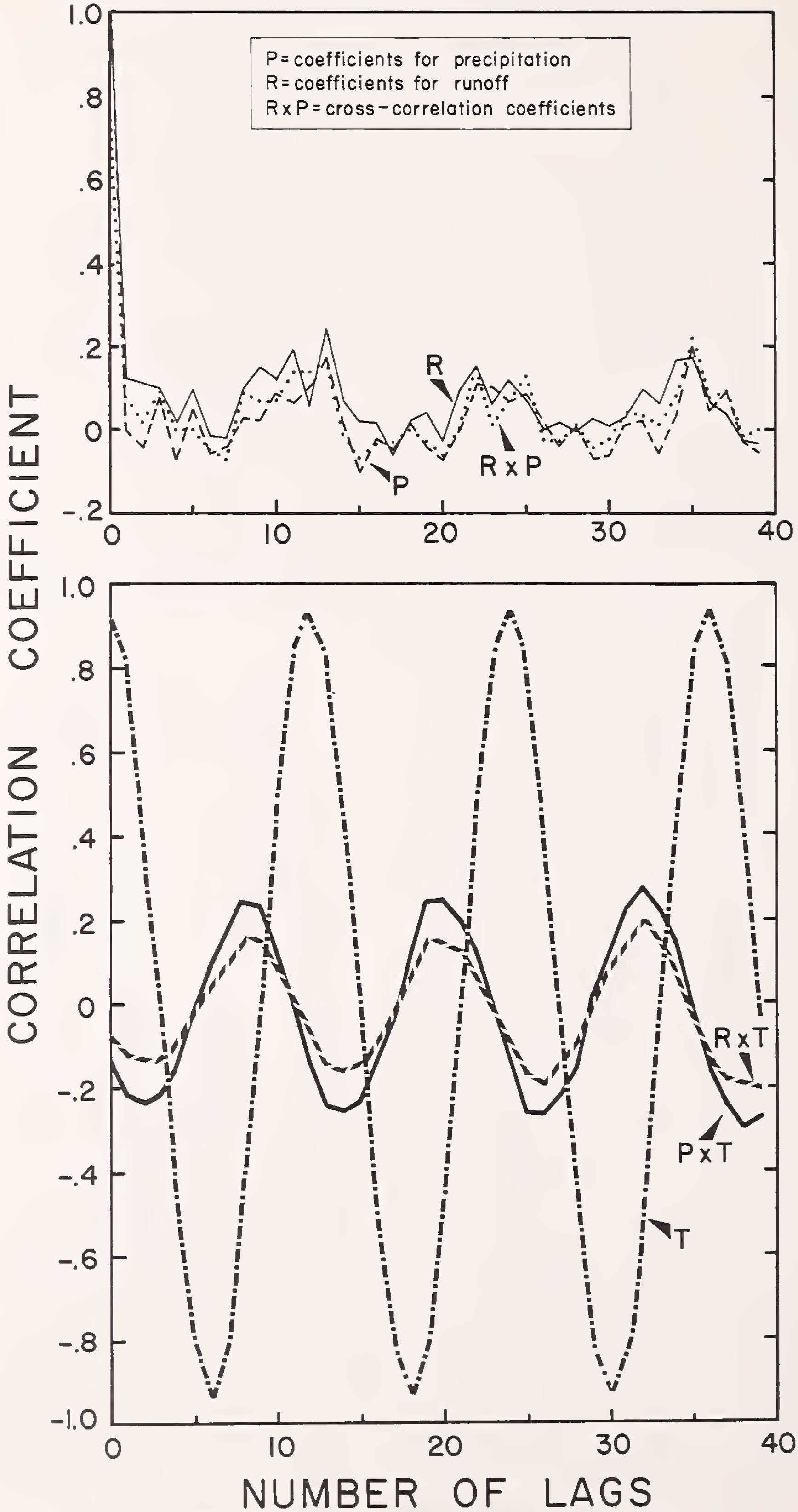

FIGURE 2.—Auto- and cross-correlation coefficients for Watkinsville watershed.

The seasonal patterns of the cross-correlation coefficients show that rainfall, runoff, and temperature are all in phase, highs of rainfall and runoff occurring with highs of temperature. This is an expression of the semitropical climate of Florida, where the summer season is the wet season of the year.

Figure 2 is a plot of the output from program I for the Watkinsville watershed. In this figure it is readily apparent that temperature follows the same strong seasonal pattern as shown in Florida. However, rainfall and runoff show only indistinct seasonal patterns. The lower part of figure 2 shows that rainfall and runoff for this watershed are not in phase with temperature. In a temperate climate under variable continental and maritime influence, the wet season tends to be in late winter and early spring.

Figures 1 and 2 indicate that both rainfall and runoff have higher serial correlations with temperature than they do with themselves or with each other. It is evident that a procedure for synthesizing a runoff series must take into account this strong seasonal pattern. Temperature could be used as an index of the seasonal pattern. However, since the autocorrelation coefficients for temperature show a seasonally repeating pattern of about 0.9, the month of the year is also a good indication of season and is much easier to use.

AUTOREGRESSION AND CROSS-REGRESSION

The information in figures 1 and 2 implies a regression of the type

$$RO_M = a + b_1 P_M + c_1 RO_{M-1} + b_2 P_{M-1} + c_2 RO_{M-2} + \ldots \quad (1)$$

In equation 1, RO_M is runoff for month M, P_M is precipitation for month M, and RO_{M-1} and P_{M-1} are runoff and precipitation for the antecedent month. Additional antecedent months can be included. The terms a, b, and c are regression coefficients. Rainfall and runoff during any month are significantly correlated, and therefore it should not be necessary to include both back rainfall and back runoff as functional (not statistical) independent variables.

Equation 1 would usually be simplified to

$$RO_M = a + b_1 P_M + C_1 RO_{M-1} + C_2 RO_{M-2} + C_3 RO_{M-3} + \ldots \quad (2)$$

Equation 2 implies an operating position for prediction as follows: A rainfall record, P_M, is available by months, and values of runoff for some number of months prior to the first month of rainfall record are also available. Runoff can then be predicted for the first month of the rainfall record. Using this synthesized runoff as back runoff with the rainfall for the second month, runoff for the second month can be generated; the process is then repeated for the full length of a real or synthetic rainfall record, P_M.

Small watersheds may be defined as those drainage areas for which soil-water processes are predominant over channel processes. For such watersheds the number of back values of runoff that can influence runoff of a succeeding month must be small. The number of back months can be estimated from graphs such as figure 1 or 2. The number is the number of lags before auto- or cross-correlation first drops to near zero. Fiering and Jackson[6] (p. 67) give a systematic test, based on the coefficient of determination, to aid in choosing the number of lags. Program II in the appendix uses only one back month, as in

$$RO_M = a + bP_M + cRO_{M-1}. \quad (3)$$

It would be a simple matter to modify this program to include additional months by using a matrix-inversion library routine.

It is usually necessary to let the final generating method evolve from systematic examination of the data set. Therefore, program II is constructed to evaluate all the regression coefficients in the set of six equations, 4 through 9.

$$RO_M = a + cRO_{M-1} \quad (4)$$
$$RO_M = a + bP_M \quad (5)$$
$$RO_M = a + bP_M + cRO_{M-1} \quad (6)$$
$$\log RO_M = a + c \log RO_{M-1} \quad (7)$$
$$\log RO_M = a + b \log P_M \quad (8)$$
$$\log RO_M = a + b \log P_M + c \log RO_{M-1} \quad (9)$$

Each of these equations is evaluated separately for each calendar month of the year; the choice of model can thus be guided by best-fit considerations as well as operational requirements.

Tables 1 through 4 show the coefficients derived by least-squares fitting of the equations to the data sets for Taylor Creek and Watkinsville. Cases A, B, and C for natural data refer respectively to equations 4, 5, and 6. Cases A, B, and C with transformed data refer to equations 7, 8, and 9.

Figure 3 is a plot of the coefficients of equation 6 for Taylor Creek. Figure 4 is a similar plot for the Watkinsville watershed. The lower part of each of these figures shows the residual errors, consisting

[6]Cited in footnote 2.

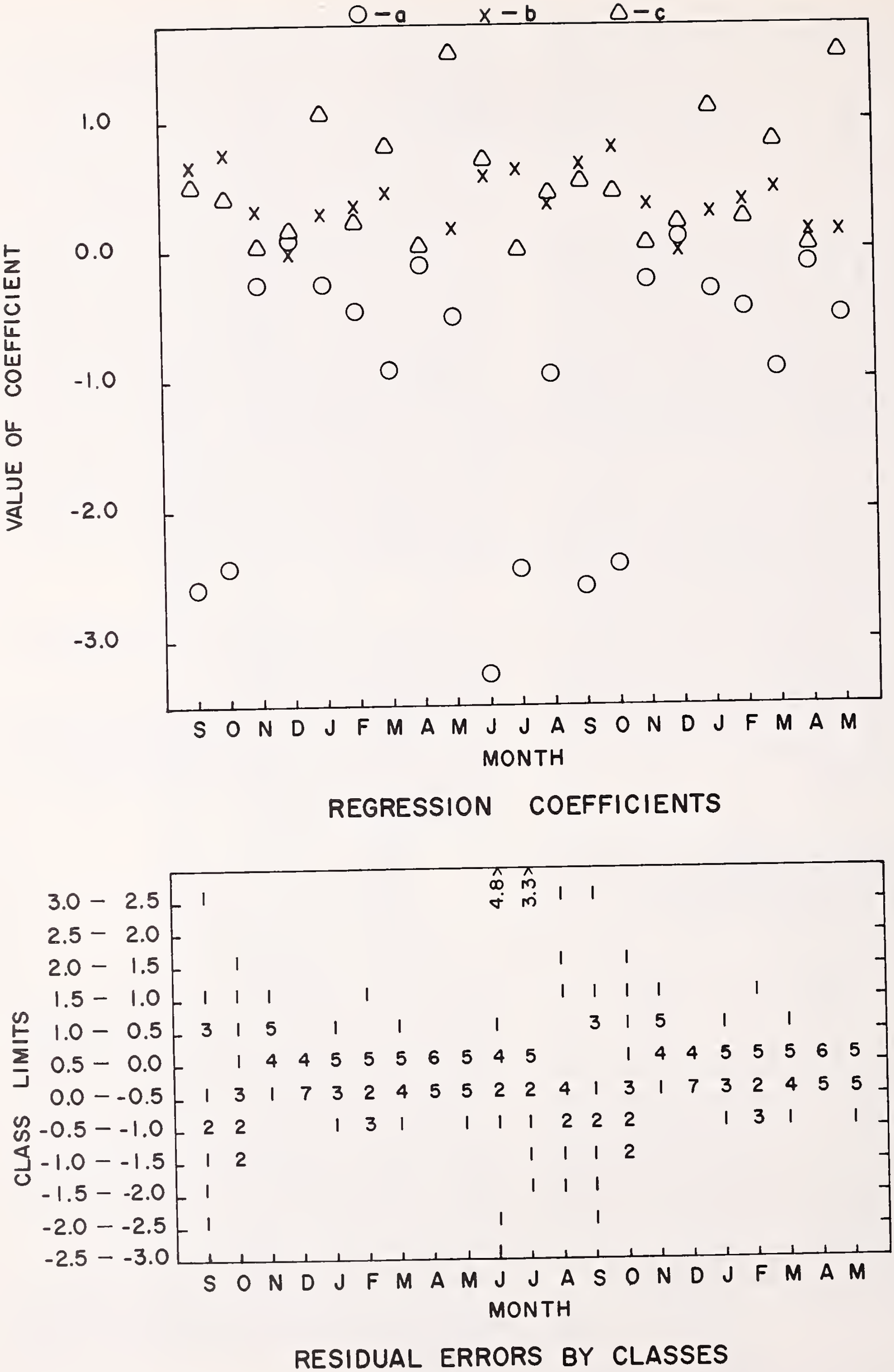

FIGURE 3.—Monthly regression coefficients and residual errors for Taylor Creek watershed.

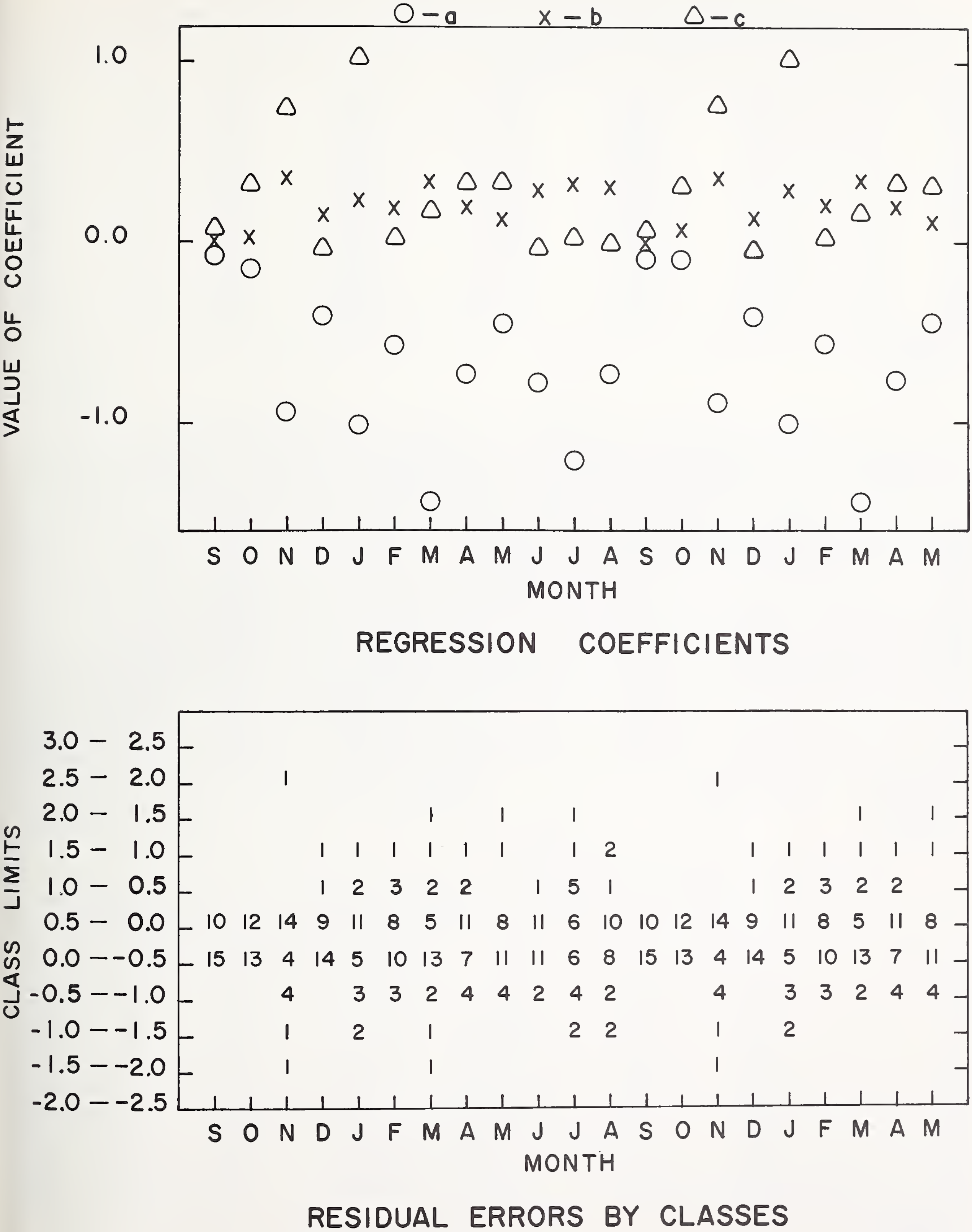

FIGURE 4.—Monthly regression coefficients and residual errors for Watkinsville watershed.

TABLE 1.—*Taylor Creek natural data: Auto- and cross-regression coefficients by month*

Month	Case A[1]		Case B[2]		Case C[3]		
	a	*c*	*a*	*b*	*a*	*b*	*c*
Jan.	0.339	0.935	−0.105	0.270	−0.307	0.276	1.101
Feb.	.481	.100	−.268	.313	−.446	.334	.279
Mar.	.387	1.022	−.625	.478	−.894	.428	.819
Apr.	.148	.078	−.034	.102	−.133	.109	.088
May	−.061	1.914	−.572	.219	−.512	.124	1.561
June	1.696	−.114	−2.771	.579	−3.308	.616	.710
July	1.474	.124	−2.411	.657	−2.407	.655	.006
Aug.	.937	.676	−1.257	.507	−.950	.352	.422
Sept.	1.000	1.067	−2.365	.802	−2.598	.676	.535
Oct.	1.531	.195	−.636	.699	−2.401	.805	.418
Nov.	.177	.061	−.060	.289	−.245	.310	.073
Dec.	.116	.149	.141	.014	.090	.167	.151

[1]Equation 4.
[2]Equation 5.
[3]Equation 6.

TABLE 2.—*Taylor Creek log-transformed data: Auto- and cross-regression coefficients by month*

Month	Case A[1]		Case B[2]		Case C[3]		
	a	*c*	*a*	*b*	*a*	*b*	*c*
Jan.	−0.359	0.547	−1.935	0.999	−0.883	0.974	0.495
Feb.	−.248	.815	−2.364	1.168	−1.067	1.015	.728
Mar.	−.089	.811	−2.831	1.536	−1.547	1.561	.824
Apr.	−1.337	.600	−2.728	.702	−1.932	.747	.605
May	−1.005	.410	−3.465	1.156	−2.575	1.145	.405
June	.263	.469	−7.190	3.335	−6.281	3.302	.446
July	.161	.643	−7.231	3.888	−5.646	3.101	.271
Aug.	−.009	1.029	−5.875	3.075	−3.221	1.732	.837
Sept.	.531	.437	−3.559	2.178	−3.005	1.917	.302
Oct.	−.440	.678	−1.069	.832	−1.528	.969	.734
Nov.	−1.678	.420	−1.752	.361	−1.682	.552	.478
Dec.	−1.207	.502	−2.127	.396	−1.316	.334	.461

[1]Equation 7.
[2]Equation 8.
[3]Equation 9.

TABLE 3.—*Watkinsville natural data: Auto- and cross-regression coefficients by month*

Month	Case A[1]		Case B[2]		Case C[3]		
	a	*c*	*a*	*b*	*a*	*b*	*c*
Jan.	0.164	1.410	−1.021	0.326	−1.014	0.272	1.065
Feb.	.444	−.061	−.482	.195	−.552	.203	.060
Mar.	.417	.522	−1.440	.349	−1.451	.337	.193
Apr.	.243	.457	−.632	.265	−.695	.233	.327
May	.053	.302	−.202	.118	−.432	.134	.331
June	.272	−.105	−.774	.281	−.771	.281	−.011
July	.439	.143	−1.210	.334	−1.214	.333	.030
Aug.	.422	−.077	−.734	.302	−.735	.303	.002
Sept.	.012	.033	.005	.007	−.036	.015	.040
Oct.	.061	.553	−.102	.069	−.107	.068	.337
Nov.	.352	−.513	−.822	.346	−.921	.358	.792
Dec.	.254	−.037	−.406	.147	−.394	.146	−.031

[1]Equation 4.
[2]Equation 5.
[3]Equation 6.

TABLE 4.—*Watkinsville log-transformed data: Auto- and cross-regression coefficients by month*

Month	Case A[1]		Case B[2]		Case C[3]		
	a	c	a	b	a	b	c
Jan.	−1.416	0.374	−7.532	3.402	−6.556	3.258	0.238
Feb.	−1.825	.183	−5.819	2.478	−5.306	2.699	.320
Mar.	−1.209	.318	−8.656	3.922	−7.914	3.863	.279
Apr.	−1.763	.242	−4.768	1.965	−4.337	1.935	.202
May	−3.228	.130	−5.057	1.364	−4.765	1.365	.131
June	−2.785	.219	−5.894	2.027	−5.163	2.014	.203
July	−.654	.669	−6.673	2.437	−4.248	2.004	.500
Aug.	−2.902	.180	−5.411	1.822	−4.952	1.802	.144
Sept.	−3.439	.187	−4.595	.532	−3.884	1.006	.338
Oct.	−3.115	.225	−4.333	.593	−3.475	.589	.210
Nov.	−4.143	−.097	−5.150	1.515	−3.957	1.815	.364
Dec.	−3.212	−.020	−7.180	2.911	−7.604	2.938	−.103

[1]Equation 7.
[2]Equation 8.
[3]Equation 9.

of the difference between the observed runoff and the predicted runoff for each month of record; they are grouped by classes. Predicted runoff is generated with the coefficients to appropriate to the calendar month, as given in the upper part of the figure.

The coefficients in figures 3 and 4 exhibit in a crude way the seasonal patterns of figures 1 and 2. The pattern is irregular because the statistical averaging in figures 3 and 4 is over only approximately one-twelfth of the data points in figures 1 and 2. For example, a lag-1 correlation coefficient is based on all pairs January–February, February–March, March–April, and so on through the years of record. The monthly regression coefficients for a calendar month, however, are based only on observations for that calendar month in the record.

SEASONALLY CONTINUOUS AUTOREGRESSION AND CROSS REGRESSION

This section describes the development of a modified form of equation 6 that is mathematically continuous in a cyclic pattern throughout the year. It should be remembered that the regression coefficients in figures 3 and 4 are in discrete sets, each calendar month producing a set. A total of 36 coefficients are needed, and must be obtained from the data by least squares. There is no continuity from month to month, and the patterns of the coefficients are quite irregular.

Equation 6 can be modified by making coefficients a, b, and c functions of calendar month. In such a modified equation, one would normally be required to estimate the parameters of the three calendar-month functions. A simpler and more direct method of establishing continuity of the coefficients is by fitting an interpolation function by least squares. The interpolating function used here was continuous parabolic interpolation[7] modified to a cyclic form.

Consider the arrangement of monthly values of the coefficient a of equation 6, shown in table 5. Base values are located as shown, one every 4 months. Continuous parabolic interpolation re-

[7]Snyder, cited in footnote 5.

TABLE 5.—*Schematic for seasonally continuous interpolation*

Month	Base value	Interpolated value
Sept.	a_9	
Oct.		
Nov.		
Dec.		
Jan.	a_1	
Feb.		a_2
Mar.		a_3
Apr.		a_4
May	a_5	
June		a_6
July		a_7
Aug.		a_8
Sept.	a_9	
Oct.		a_{10}
Nov.		a_{11}
Dec.		a_{12}
Jan.	a_1	
Feb.		
Mar.		
Apr.		
May	a_5	

quires four base values for an interpolated value. The values for a_2, a_3, and a_4 are based on a_9, a_1, a_5, and a_9. It can be seen that a seasonally continuous cyclic system of interpolation is produced.

Program III, listed in the appendix, was written to derive, simultaneously by least squares, the base values a_1, a_5, a_9, b_1, b_5, b_9, c_1, c_5, and c_9. The squares of errors for every month of the record are utilized, rather than just the errors for those months where the base values are located. It can be noted in the program listing that interpolation is carried out using numerical interpolating coefficients. These coefficients were taken from table 7 in Snyder.[8]

Tables 6 and 7 are listings of the base values, defined in table 5, for the three seasonal functions for the coefficients of equation 6. Several different determinations of these seasonal functions were made, using different fitting criteria. The criteria for and deviations of coefficients are discussed below.

The coefficients for natural data, unrestrained, in tables 6 and 7 are the seasonally continuous versions of the discrete sets of coefficients in figures 3 and 4. In further explanation, in table 8 the coefficients for natural data, unrestrained, are expanded to include all calendar months, using the same interpolating coefficients as in the listing of program III. These expanded coefficients are plotted as continuous lines in figure 5 for both Taylor Creek and Watkinsville. For convenient comparison, the discrete coefficients from figures 3 and 4 are also plotted. It can be seen that the cyclically continuous coefficients form a smooth average of the discretely derived coefficients. By evaluating only 9 parameters from the historical record, instead of the 36 in monthly discrete analysis, a greater degree of averaging is imposed. This averaging tends to smooth the irregularities in the discrete coefficients.

The comparisons evident in figure 5 are based on equation 6 as the predictor model. Similar comparisons could be made for all other models implied in equations 4 through 9. The researcher may need to make such additional comparisons before choosing the predictor model for a specific watershed.

Before discussing some of the additional fitting criteria in tables 6 and 7, it is necessary to consider the plots in figure 6. The upper plot shows the position a line would take if fitted by customary least-squares procedures. In particular, the positive error (observed minus calculated) for data point 1 would be roughly balanced by the negative error for point 3. If this fitted line is subsequently used for simulating values of Y based on values of X, a bias results. To the left of point 2, calculated values of Y are negative. These can be set to zero, since negative runoff does not occur. However, from point 2 to point 3, calculated values of runoff are positive, whereas in reality they should also be zero. Since the positive values from point 2 to point 3 are not

[8]Snyder, W. M. 1962. Closure to continuous parabolic interpolation. J. Hydraul. Div., Proc. Am. Soc. Civ. Eng. 88 (HY4): 265–274.

TABLE 6.—*Taylor Creek seasonally cyclic coefficients in* $RO_M=a_M+b_MP_M+c_MRO_{M-1}$

Coefficient	Fitting criteria: ([1])	([2])	Natural $RO\geq 0$	Natural $b\leq 1$ $RO\geq 0$	Estimated natural for simulation
a_1	−.6763	−0.1540	−12.1178	−24.7986	−3.50
a_5	−3.0677	−1.0969	−3.6173	−4.9686	−1.71
a_9	−2.5040	−2.1945	−2.6206	−3.2856	−2.34
b_1	.7140	.3653	2.2952	1.0000	.85
b_5	1.7859	.3411	.7240	.8315	.46
b_9	1.5078	.6773	.6865	.7178	.68
c_1	.6396	.2367	1.2124	2.0082	.49
c_5	.5463	.0302	.1887	.1050	.07
c_9	.5400	.3246	.4268	.5433	.35
S_r	.8819	1.0965	.9487	1.5490	1.05
$\hat{Y}$	−1.1115	1.1455	1.0458	1.1709	1.1439
$\bar{Y}$	−1.1115	1.1439	1.1439	1.1439	——

[1]Log-transformation of data—unrestrained.
[2]Natural data—unrestrained.

S_r Standard deviation of residual errors.

$\hat{Y}$ Mean of calculated values.
$\bar{Y}$ Mean of observed values.

TABLE 7.—*Watkinsville seasonally cyclic coefficients in* $RO_M = a_M + b_M P_M + c_M RO_{M-1}$

Coefficient	Fitting criteria		
	(1)	(2)	Natural $RO>0$
a_1	−5.5750	−0.8104	−3.1206
a_5	−5.0523	−.8910	−2.0072
a_9	−3.6056	−.5464	−5.0934
b_1	2.5766	.2556	.5482
b_5	2.1564	.2684	.4032
b_9	.7799	.2355	.7652
c_1	.2144	.1165	.2612
c_5	.2632	.2769	.4624
c_9	.2432	−.0124	.1258
S_e	1.2789	.5868	.4022
$\hat{Y}$	−3.1350	.3380	.3104
$\bar{Y}$	−3.1350	.3380	.3380

[1]Log-transformation of data—unrestrained.
[2]Natural data—unrestrained.

S_e Standard deviation of residual errors.
$\hat{Y}$ Mean of calculated values.
$\bar{Y}$ Mean of observed values.

balanced by negative values from point 1 to point 2, the average of the simulated values is higher than the average of the observed values.

TABLE 8.—*Examples of complete sets of seasonally cyclic coefficients*

Watershed and month	Coefficient a	b	c
Taylor Creek:			
Jan.	−0.1540	0.3653	0.2367
Feb.	−.0222	.3306	.1817
Mar.	−.4293	.3127	.1095
Apr.	−.7804	.3151	.0493
May	−1.0969	.3411	.0302
June	−1.4340	.4150	.0775
July	−1.8322	.5272	.1700
Aug.	−2.1372	.6304	.2661
Sept.	−2.1945	.6773	.3246
Oct.	−1.8351	.6381	.3323
Nov.	−1.1839	.5438	.3120
Dec.	−.5279	.4383	.2760
Watkinsville:			
Jan.	−0.8104	0.2556	0.1165
Feb.	−.8534	.2604	.1675
Mar.	−.8887	.2653	.2228
Apr.	−.9050	.2686	.2676
May	−.8910	.2684	.2769
June	−.8204	.2621	.2018
July	−.7072	.2515	.1342
Aug.	−.5997	.2410	.0410
Sept.	−.5464	.2355	−.0124
Oct.	−.5739	.2370	−.0103
Nov.	−.6518	.2427	.0239
Dec.	−.7430	.2499	.0723

A modification of the fitting procedure should significantly reduce this bias. In the lower plot in figure 6 is shown the position a line will take if negative calculated values are set to zero during fitting of the line. Errors at points 1 and 2, and perhaps at point 3, become zero. The line takes the position shown, minimizing error for points 3, 4, and 5. This procedure is not the same as deleting points with zero runoff from the data set prior to fitting. It is not known ahead of time whether point 3 would have a negative or positive calculated value; consequently, it is not known ahead of time whether point 3 should be deleted or retained. In a real data set many points could be expected to lie near the critical value of point 3.

The line shown in the lower plot of figure 6 cannot be fitted by customary least squares. It can, however, be routinely fitted by the iterative procedures of nonlinear least squares. The method is contained in program III in the appendix. The coefficients for natural data, under the restraint that predicted runoff cannot be negative, are shown in the third column of tables 6 and 7.

Table 7 shows that for Watkinsville the incorporation of the runoff restraint tended to increase the slopes of the two lines (increase in coefficient b and in coefficient c) and to decrease significantly the value of the intercept coefficient a. This follows the results expected from the sketches in figure 6. Table 6 shows that this single restraint did not produce satisfactory results for Taylor Creek. Coefficient b_1 is 2.2952. Physically, this value may not

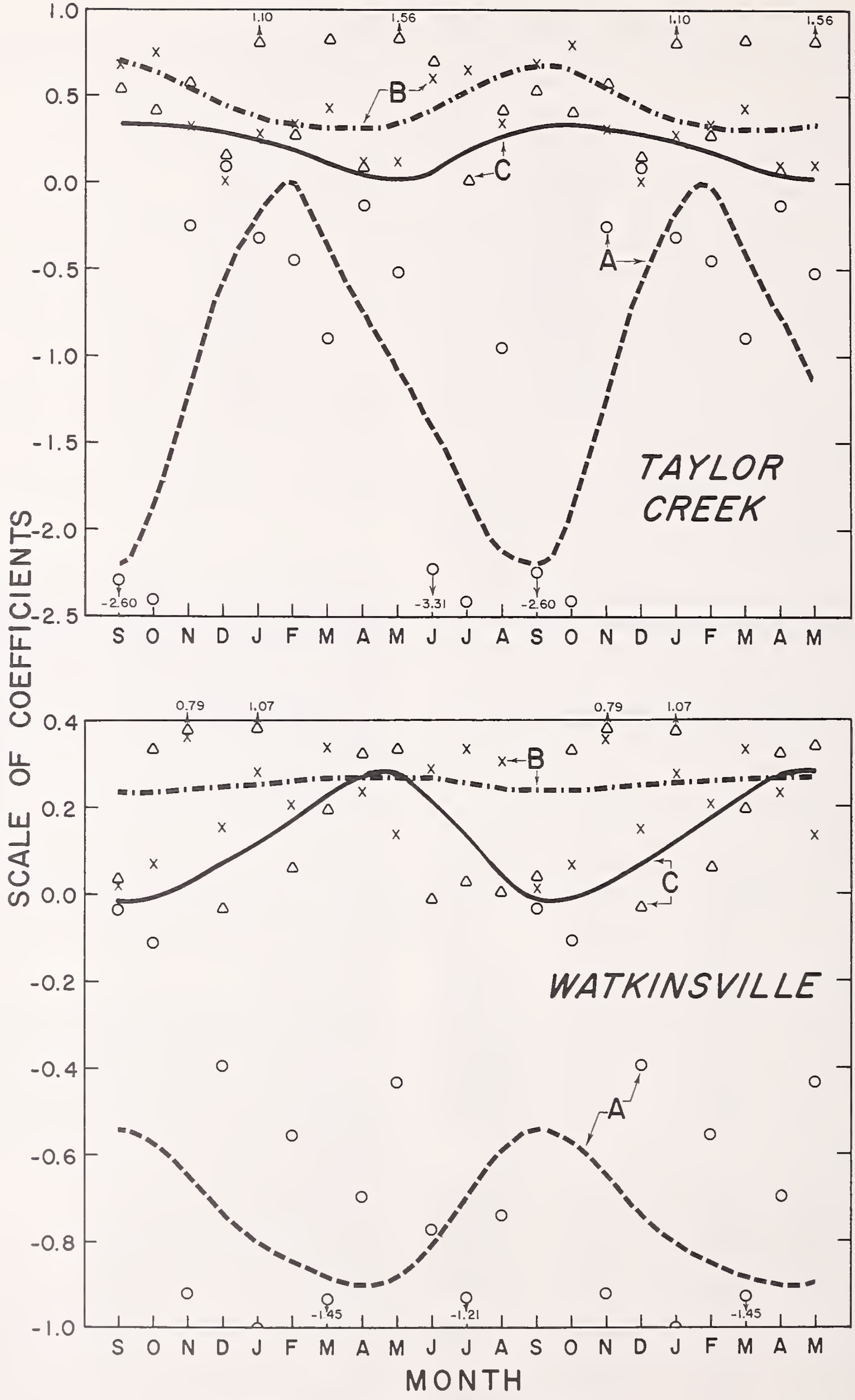

FIGURE 5.—Seasonally continuous regression coefficients.

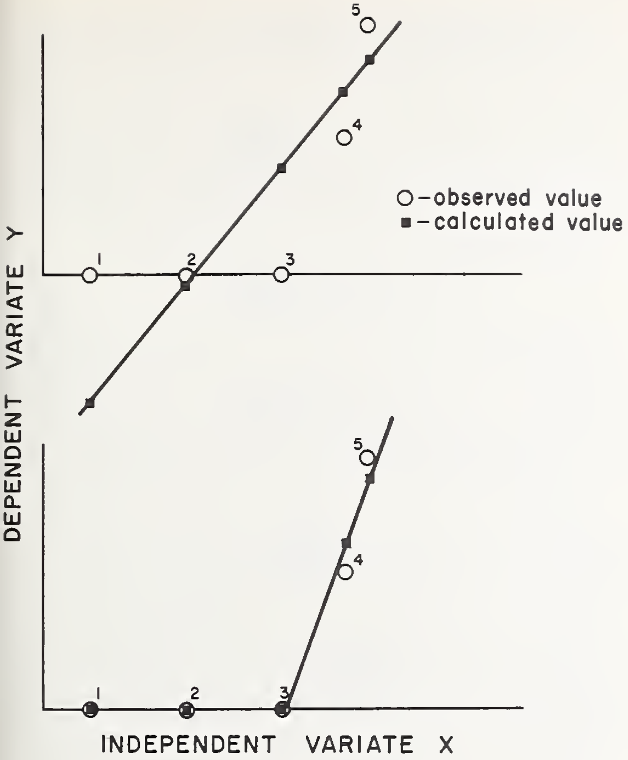

FIGURE 6.—Effect of setting negative errors to zero.

be more than 1.0; otherwise runoff at some value of rainfall would become greater than the rainfall. The Taylor Creek data were rerun under the additional restraint that the coefficients b could not be greater than 1.0. The results in table 6 under this fitting criterion show that this procedure is numerically feasible. However, the change in the coefficients is quite large. The distortion is due to some extreme values of rainfall occurring in the short record of 11 years. It is expected that a longer record would produce better averaging.

The last column in table 6 shows an estimated set of coefficients used for simulating a runoff record for Taylor Creek. These will be treated further in a later section.

SEASONALLY CONTINUOUS STATISTICAL FREQUENCY DISTRIBUTIONS

In order to use equation 6 as the predictor model for generating a series of future runoff events, it is necessary first to generate values of synthetic monthly rainfall. The usual procedure is to generate random numbers and then scale and transform these numbers to represent stochastic monthly rainfalls. Snyder and Wallace[9] developed the three-parameter log-normal distribution as a functional variate transformation of an embedded normal distribution. The variance of the embedded normal distribution was standardized at unity. The mean of the embedded normal and two parameters in the transformation function were evaluated by nonlinear least squares. When so defined and evaluated, the three-parameter log-normal distribution is a good device for generating stochastic rainfall values. One first generates pseudorandom normal numbers of a computer. These have zero mean and unit variance. They are transformed to variates of the embedded normal distribution by simple addition of the parametric mean of this embedded distribution. These variates are then transformed by the variate transformation function, after which they can be considered synthetic rainfall values.

The log-normal probability-density function used by Snyder and Wallace is

$$p(v) = [\sqrt{2\Pi}\, k\, (v-o)]^{-1} \exp\left\{-\tfrac{1}{2}\left[\frac{\ln (v-o)}{k} - m\right]^2\right\}. \qquad (10)$$

The variate transformation function is

$$\ln (v-o) = kx. \qquad (11)$$

In equations 10 and 11, x is the variate of the embedded normal distribution of unit variance and mean m. The variate v, to be regarded here as a synthetic rainfall value, is dependent on x by equation 11, utilizing the mathematical parameters o and k. Snyder and Wallace evaluated the three parameters o, k, and m by fitting equation 10 to a histogram of a rainfall record using nonlinear least squares.

The usual procedure in evaluation of statistical frequency functions such as equation 10 is analogous to the procedure for monthly discrete regression coefficients discussed earlier. An advance in procedure discussed below is the development of seasonally continuous frequency distributions analogous to the seasonally continuous regression coefficients.

In the preceding section, values of the coefficients a, b, and c were made seasonally continuous by fitting an interpolating function. The same device was used to make the statistical frequency distributions seasonally continuous. Parameters o and k in equation 10 were defined by interpolating functions, using the same arrangement as in table 5. The

[9]Cited in footnote 4.

parameter m was specified as having the same value for all months. This says that the embedded normal is the same for all months. Now in simulation, one generates stochastic values of the embedded normal as outlined above, and then reflects these to the proper calendar month using appropriately interpolated values of o and k in equation 11.

An increase in efficiency of utilization of historical data results from this seasonally continuous stochastic model. If equation 10 were fitted to the rainfall record for each calendar month, 36 parameters would be evaluated. By seasonal interpolation for o and k, only seven parameters are evaluated, roughly a 5:1 gain in efficiency. The numerical effect, as with the seasonally continuous regression coefficients, is to produce better averaging and a smoothing of irregularities.

The procedure may be viewed as fitting 12 statistical frequency distributions simultaneously through nonlinear least squares.

Program IV in the appendix is a subprogram written to fit the 12 frequency distributions simultaneously. The monthly rainfall data for Taylor Creek and Watkinsville were organized into histograms by months as shown in figure 7. The seven parameters of the seasonally continuous version of equation 10 were then evaluated. The specific distributions for each month were calculated and superimposed on the historical histograms in figure 7. The seasonal left-right shifting of the frequency functions is due to seasonal continuity of parameter o. The seasonal change in peakedness is due to seasonal continuity of parameter k.

The reader should fully understand the fitting objective in figure 7. It will be noted that the histograms have a base of 16 classes. In fitting, class errors are calculated for $12 \times 16 = 192$ classes. The sum of squares is minimized for the 192 classes. Figure 7 does not illustrate a best fit for any particular month, but an overall best fit of 12 months. The 16-class base was used because this was, by coincidence, the width required for both watersheds, for September in Taylor Creek and for July and November at Watkinsville. The concept of empty classes beyond the range of historical data,[10] thus applies to all months except these three.

Numerical values of the parameters derived from the historical data are listed in table 9. Two sets of parameters are shown for each drainage area. The origin of the weighting values in the table is given below.

[10]Snyder, cited in footnote 4.

Grant[11] investigated the theoretical basis of fitting distribution functions by nonlinear least squares. Two findings are important here. The method is identical to maximum-likelihood estimation for classified data if the class errors are normally distributed. A bias in estimation can be reduced by using weighted nonlinear least squares.

The weighting of errors used by Grant is

$$e_{wc} = [h_c - n(v)_c]/[n(v)_c]^{\gamma}. \qquad (12)$$

In equation 12, the usual statistical error is the observed number in the class h_c, minus the average value of the calculated distribution function for the class $n(v)_c$. This error is weighted by division by $n(v)_c$ exponentially scaled by the parameter γ. Grant pointed out that if γ is zero, the weighted errors for a class, e_{wc}, become identically the unweighted errors. If γ is unity, the method of nonlinear least squares becomes identical with the method of minimum chi-square. Grant found that a value for γ of 0.75 minimized the bias in estimation of the parameters in the two-parameter gamma distribution. The values of γ best suited to other distributions is apparently completely unexplored.

The two sets of parameters in table 9 are the results of fitting by weighted nonlinear least squares, with γ set to 0.50 and 0.75. The two sets of parameters for the two watersheds in table 9 are different, but the only relatively large difference is in o_5 for Watkinsville. If equation 12 is solved for v,

[11]Grant, J. L. 1973. Statistical frequency analysis by optimization of density functions to histograms. 176 pp. Doctoral dissertation, Georgia Institute of Technology, Atlanta.

TABLE 9.—*Parameters of seasonally cyclic distribution functions*

Watershed and parameter	Weight	
	0.50	0.75
Taylor Creek:		
o_1	−1.5545	−1.6005
o_5	.7175	.7300
o_9	.3378	.4530
k_1	.4953	.5250
k_5	.7976	.7663
k_9	.8097	.8042
m	1.8066	1.8176
Watkinsville:		
o_1	.5970	.5296
o_5	.3909	.1981
o_9	−.5786	−.6066
k_1	.7124	.6903
k_5	.7695	.7440
k_9	.6228	.6130
m	1.7534	1.8159

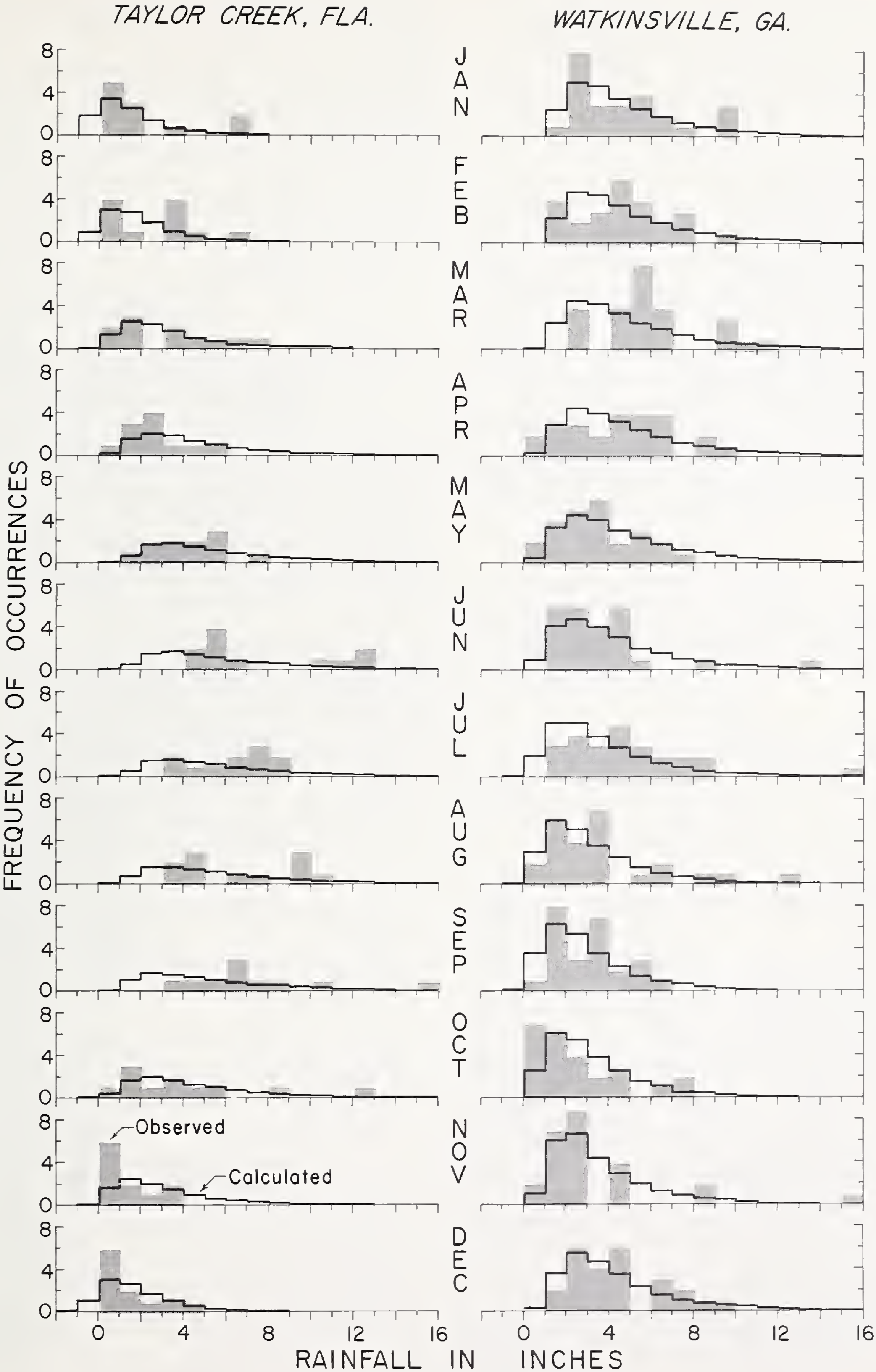

FIGURE 7.—Seasonally continuous rainfall distributions.

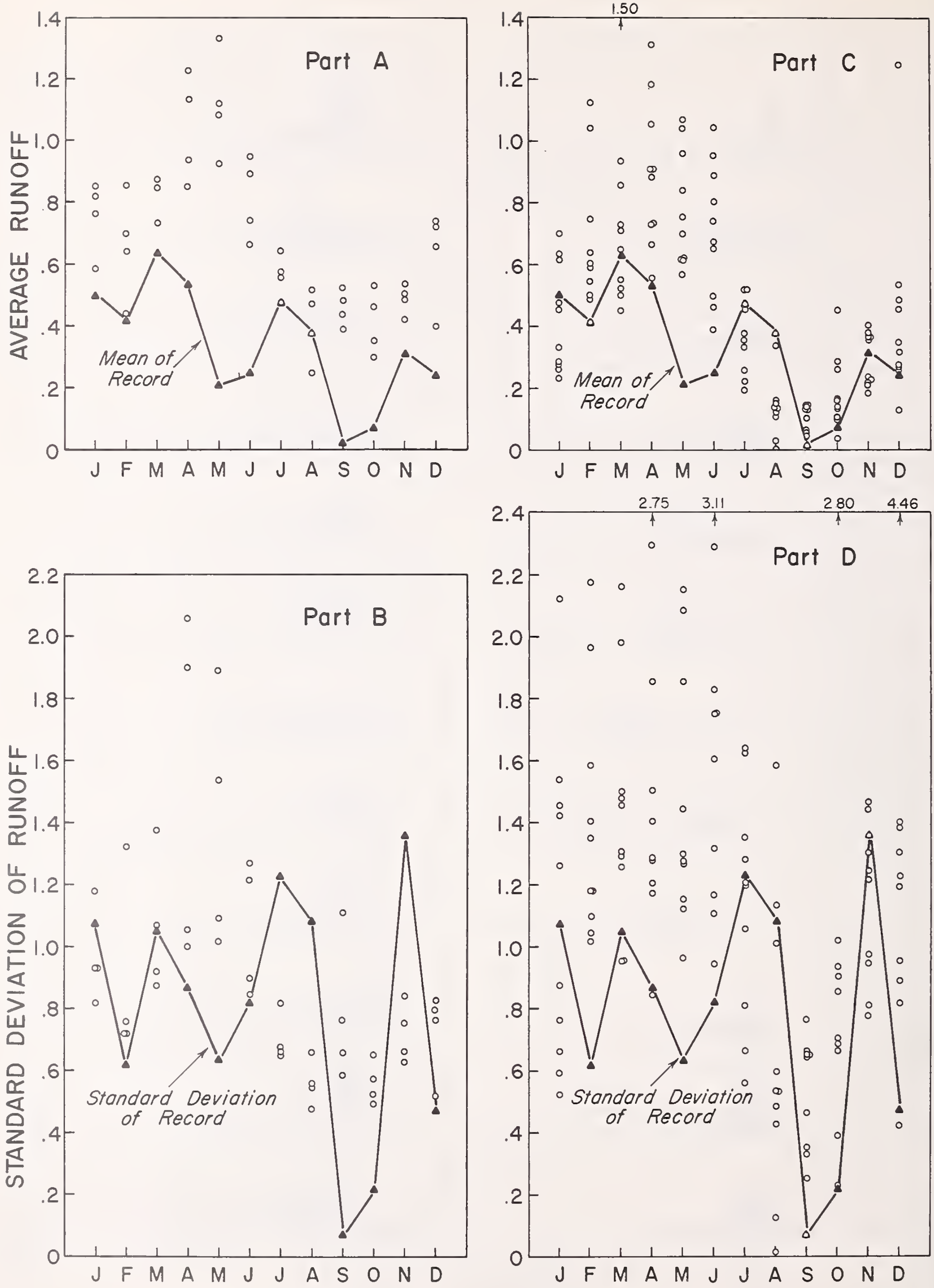

FIGURE 8.—Runoff simulation comparisons for Watkinsville watershed.

the variate transformation function takes the form

$$v = \exp(k_M x) + o_M. \tag{13}$$

The two different values for o_5 in table 9 thus take on the physical significance that a simulated rainfall value is changed by about 0.2 inch. This does not seem of major significance for a stochastic variate in an observed range from 0 to 16 inches. It is, however, an example of the detailed exploration necessary in further development of weighted nonlinear least squares.

The parameter values obtained with $\gamma = 0.75$ were used in the following section.

GENERATION OF SYNTHETIC RUNOFF SERIES

Programs I through IV, discussed in previous sections, are essentially analytical. That is, they are intended for analysis of historical data, to extract information in the form of numerical values of coefficients and parameters. Program V is not a data-analysis program; it brings together elements from the preceding programs to synthesize new information.

The basic synthesizing mechanism in program V has equation 6 as its core. However, researchers can easily modify this program if it appears desirable, after processing a particular historical record through programs I through IV, to use another of generating equations 4 through 9.

The full generating mechanism for synthesis requires adding an error to equation 6, as in—

$$RO_M = a_M + b_M P_M + cRO_{M-1} + e. \tag{14}$$

This error is implicit, of course, in the very definition of least squares. It is the error residual to fitting a set of equations to a historical data set. Such errors are usually assumed to be normally distributed.

In using equation 14 to generate a synthetic data series, both e and rainfall P_M must be considered stochastic variates. They are calculated from computer-generated pseudorandom numbers, by the following method.

Normally distributed variates may be standardized to zero mean and unit standard deviation. Standardization is performed by subtracting the mean from each item in a sample and dividing the difference by the standard deviation. Since any two standardized variates must be equal, one can write immediately

$$\frac{r - m_r}{s_r} = \frac{e - m_e}{s_e}. \tag{15}$$

In equation 15, r is the pseudorandom number, m_r is the mean of r, s_r is the standard deviation of r, e is the stochastic variate of equation 14, m_e is the mean of e, and s_e is the standard deviation of e. If m_r is zero and s_r is unity, then e is expressed

$$e = rs_e + m_e. \tag{16}$$

Equation 16 converts a computer-generated pseudorandom number r to the stochastic variate e.

Theoretically, m_e is zero when the method of least squares is employed. For nonlinear least squares m_e should also be zero. In practice it is sometimes very small, but not zero. This nonzero m_e is the difference between the means of calculated values and means of observed values in tables 6 and 7.

The rainfall stochastic variate is generated from the continuous seasonal distribution of the preceding section. It should be recalled that the distribution is based on an embedded normal distribution. The standardizing equation is

$$\frac{r - m_r}{s_r} = \frac{x - m_x}{s_x}. \tag{17}$$

Now, if m_r is zero and s_r is unity as before, and s_x, the standard deviation of the embedded normal, is unity by definition, then

$$x = r + m_x. \tag{18}$$

Substituting equation 18 into the variate transformation equation and specifically using P for precipitation in place of v gives

$$P_M = \exp\,[k_M(r + m)] + o_M. \tag{19}$$

With e and P_M generated from pseudorandom numbers r, using equations 16 and 19, equation 14 is readily evaluated.

Figure 8 illustrates the results of runoff simulation for Watkinsville. The monthly means of generated data are shown at the top, and standard deviations are shown at the bottom. To the left are results based on coefficients derived by using natural data, unrestrained, taken from table 7. To the right are results based on coefficients derived by using natural data with calculated negative runoff values set to zero during fitting, also from table 7. The positive bias in synthetic data, illustrated by the upper plot in figure 6, is clearly shown in part A of figure 8. The generated means of 50 years for four sets lie above the means of record. In part C of figure 8 are the means of 10 sets of 50 years each, using coefficients resulting from restrained fitting. Here the means have been lowered toward the mean of record.

It is not possible to see the same effect in the

standard deviations in parts B and D of figure 8. The standard deviations were lowered for the months of August, September, and October, but not appreciably for other months.

Fiering and Jackson[12] state that first and second moments should be preserved during simulation. This requirement is not met in figure 8, possibly because of the relative lengths of real and simulated records. Fontane[13] found somewhat similar results with respect to sampling the third moment of a gamma distribution.

The effect of record length on distributions bounded on one side may be nonrigorously described as follows. Consider an infinitely long rainfall record. Within this record, the items are bounded by a low value of zero. Scattered through the record are a few very large values. Now consider a short time span and a long time span to be selected from the infinitely long record. The chance of including a very large event is greater for the long sample than for the short sample. The same reasoning indicates that a long simulated record is more likely to contain a very large event than a short real record.

If the distributions were unbounded, one could expect extreme positive values to balance extreme negative values as sample size increased. No such balancing is possible with a distribution bounded on one side. The effect is shown graphically in figure 9.

In figure 9, histograms of rainfall are compared for three different calendar months. The observed record length is 25 years. The simulated record length is 50 years. In February, the observed highest monthly rainfall was between 9 and 10 inches. For the simulated record the highest value doubled, to 20–21 inches. For June the highest simulated value is also nearly double the observed value. These extreme values exert tremendous influence on the moments of the samples. How much of this effect is present in figure 8 cannot be answered at this time, because it cannot be known whether the seasonal log-normal distribution is in fact the proper distribution.

[12]Cited in footnote 2.

[13]Fontane, D. G. 1970. Statistical tolerance limits for a Pearson type III distribution. 104 pp. Master's thesis, Georgia Institute of Technology, Atlanta.

TABLE 10.—*Results of simulation for Taylor Creek watershed*

Month	Observed	200-year simulation (1)	(2)	(3)	(4)
		Mean runoff			
Jan.	0.451	0.786	0.506	0.008	0.101
Feb.	.526	.724	.268	0	.055
Mar.	.926	.786	.083	0	.311
Apr.	.220	1.030	.106	0	.895
May	.360	1.428	1.182	.130	1.466
June	1.655	1.467	1.521	2.774	1.842
July	1.679	2.676	1.971	4.086	2.710
Aug.	2.072	2.991	2.614	4.426	3.006
Sept.	3.210	2.936	3.340	3.771	3.091
Oct.	2.157	2.810	2.840	1.699	2.710
Nov.	.308	1.584	2.440	.381	1.342
Dec.	.162	1.031	1.192	.058	.952
		Standard deviation of runoff			
Jan.	0.725	0.906	3.148	0.117	0.484
Feb.	.828	.944	2.146	0	.255
Mar.	1.350	1.124	.986	0	1.101
Apr.	.270	1.143	.687	0	1.729
May	.650	1.760	3.083	1.467	1.896
June	2.719	1.820	2.670	4.253	2.245
July	1.807	3.065	3.241	4.479	4.208
Aug.	1.985	3.864	3.192	4.064	3.575
Sept.	3.384	3.716	4.041	4.212	4.639
Oct.	2.971	3.235	4.207	3.441	3.949
Nov.	.539	1.943	5.091	1.812	2.193
Dec.	.135	1.193	4.125	.814	.398

[1]Natural data—unrestrained.

[2]$RO \geq 0$ required during coefficient evaluation.

[3]$RO \geq 0$ and $b \leq 1$ required during coefficient evaluation.

[4]Estimated coefficients (table 6).

The uncertainty described above is caused in part by using one short record. Similar uncertainty resulted when runoff was simulated for Taylor Creek.

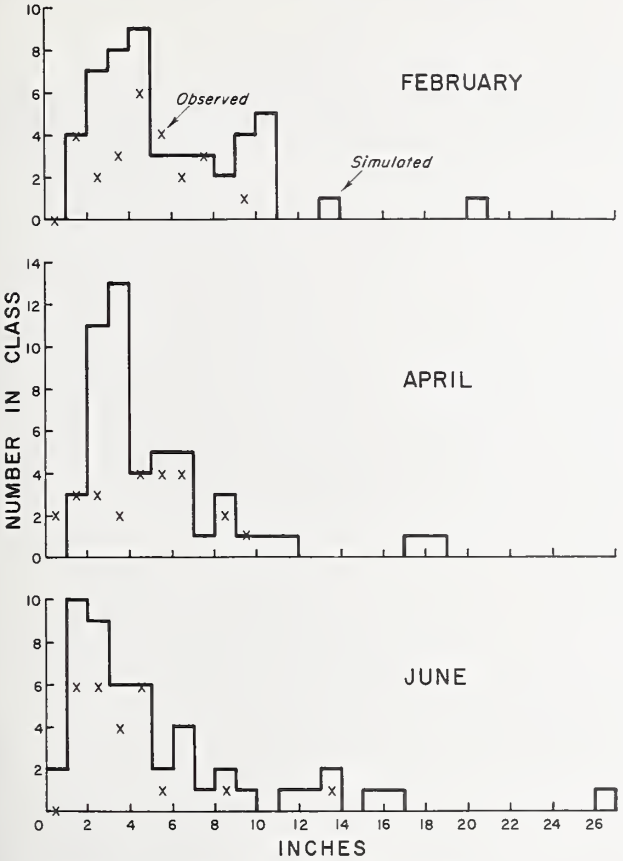

FIGURE 9.—Simulated and recorded rainfall distributions for Watkinsville watershed.

The resulting average values of means and standard deviations are shown in table 10. Four different 200-year simulations were performed corresponding to the four sets of coefficients derived from untransformed data and listed in table 6.

Not one of the four simulation results in table 10 is very close to the record averages of means and standard deviations, but the record length of real data was only 11 years. One is hard pressed to decide whether the simulated values are rational future expectations. Within the criteria specified they are valid statistical expectations.

CONCLUSIONS

Synthetic data such as generated by this package of five computer programs for analyzing and synthesizing time series data are needed for use with watershed models, to predict possible runoff of water, sediment, and chemicals.

The illustrations do not represent a fully tested set of procedures, but are intended, rather, to show the difficulties met when a synthesis is based on the very short record available for the average research watershed.

We hope that other researchers will be interested in applying these procedures, or modified versions of them, to simulation problems on other watersheds. Pooling of results should make possible much better tests and refinements of procedures than could be obtained with the limited resources at one location.

APPENDIX.—COMPUTER PROGRAMS

Program I

Input variable definitions

NSET	Number of data sets to be run.
N	Number of observations in a data set.
LAG	Number of lags for which auto- and cross-correlation coefficients are to be computed. LAG=1 is zero lag.
RO	Values of the dependent variable (runoff).
PR	Values of the independent variable (precipitation).

Output variable definitions

SRO1	Sum of the lagged dependent values.
SSR1	Sum of squares of SRO1.
SRO2	Sum of the unlagged dependent values.
SSR2	Sum of squares of SRO2.
SPR1	Sum of the lagged independent values.
SSP1	Sum of squares of SPR1.
SPR2	Sum of the unlagged independent values.
SSP2	Sum of squares of SPR2.
SSRO	Sum of products of SRO1 with SRO2.

SSPR	Sum of products of SPR1 with SPR2.	CORPR	Autocorrelation coefficients for independent variable.
SPPR	Sum of products of SRO1 with SPR2.	CPRRO	Coefficients of cross correlation of variables.
CORRO	Autocorrelation coefficients for dependent variable.		

```
C COMPUTATION OF AUTO- AND CROSS-CORRELATION COEFFICIENTS
      DIMENSION RO(1000),PR(1000),SRO1(40),SRO2(40),SSR1(40),SSR2(40),SP
     1R1(40),SPR2(40),SSP1(40),SSP2(40),SSRO(40),SSPR(40),SPPR(40),CORRO
     1(40),CORPR(40),CPRRO(40)
      READ(5,100) NSET
      DO 111 LSET=1,NSET
      READ(5,100) N,LAG
  100 FORMAT(2I5)
      READ(5,101) (RO(I),I=1,N)
      READ(5,101)(PR(I),I=1,N)
  101 FORMAT(20F4.2)
      DO 104 I=1,LAG
      SRO1(I)=0.0
      SRO2(I)=0.0
      SSR1(I)=0.0
      SSR2(I)=0.0
      SPR1(I)=0.0
      SPR2(I)=0.0
      SSP1(I)=0.0
      SSP2(I)=0.0
      SSRO(I)=0.0
      SSPR(I)=0.0
  104 SPPR(I)=0.0
      DO 106 I=1,LAG
      NEND=N-I+1
      DO 106 J=1,NEND
      SRO1(I)=SRO1(I)+RO(I+J-1)
      SSR1(I)=SSR1(I)+RO(I+J-1)*RO(I+J-1)
      SRO2(I)=SRO2(I)+RO(J)
      SSR2(I)=SSR2(I)+RO(J)*RO(J)
      SPR1(I)=SPR1(I)+PR(I+J-1)
      SSP1(I)=SSP1(I)+PR(I+J-1)*PR(I+J-1)
      SPR2(I)=SPR2(I)+PR(J)
      SSP2(I)=SSP2(I)+PR(J)*PR(J)
      SSRO(I)=SSRO(I)+RO(I+J-1)*RO(J)
      SSPR(I)=SSPR(I)+PR(I+J-1)*PR(J)
  106 SPPR(I)=SPPR(I)+RO(I+J-1)*PR(J)
      WRITE(6,112)
  112 FORMAT(*0 SUMS AND SUMS OF PRODUCTS*)
      WRITE(6,108) (I,SRO1(I),I=1,LAG)
      WRITE(6,108) (I,SSR1(I),I=1,LAG)
      WRITE(6,108) (I,SRO2(I),I=1,LAG)
      WRITE(6,108) (I,SSR2(I),I=1,LAG)
      WRITE(6,108) (I,SPR1(I),I=1,LAG)
      WRITE(6,108) (I,SSP1(I),I=1,LAG)
      WRITE(6,108) (I,SPR2(I),I=1,LAG)
      WRITE(6,108) (I,SSP2(I),I=1,LAG)
      WRITE(6,108) (I,SSRO(I),I=1,LAG)
      WRITE(6,108) (I,SSPR(I),I=1,LAG)
      WRITE(6,108) (I,SPPR(I),I=1,LAG)
  108 FORMAT(*0*, 6(I5,E16.8))
      DO 107 I=1,LAG
```

```
    NMI=N-I+1
    CORRO(I)=(SSRO(I)-(1.0/(NMI))*SRO1(I)*SRO2(I))/SQRT((SSR2(I)-(1.0/
   1(NMI))*SRO2(I)*SRO2(I))*(SSR1(I)-(1.0/(NMI))*SRO1(I)*SRO1(I)))
    CORPR(I)=(SSPR(I)-(1.0/(NMI))*SPR1(I)*SPR2(I))/SQRT((SSP2(I)-(1.0/
   1(NMI))*SPR2(I)*SPR2(I))*(SSP1(I)-(1.0/(NMI))*SPR1(I)*SPR1(I)))
    CPRRO(I)=(SPPR(I)-(1.0/(NMI))*SPR2(I)*SRO1(I))/SQRT((SSP2(I)-(1.0/
   1(NMI))*SPR2(I)*SPR2(I))*(SSR1(I)-(1.0/(NMI))*SRO1(I)*SRO1(I)))
107 CONTINUE
    WRITE(6,109)
109 FORMAT(*0  AUTO-CORRELATION OF DEPENDENT VARIABLE*)
    WRITE(6,105)(I,CORRO(I),I=1,LAG)
105 FORMAT(*0*,10(I5,F7.4))
    WRITE(6,110)
110 FORMAT(*0 AUTO-CORRELATION OF INDEPENDENT VARIABLE*)
    WRITE(6,105)(I,CORPR(I),I=1,LAG)
    WRITE(6,113)
113 FORMAT(*0 CROSS-CORRELATION OF VARIABLES*)
    WRITE(6,105)(I,CPRRO(I),I=1,LAG)
111 CONTINUE
    END
```

Program II

Input variable definitions

NSETS	Number of data sets to be run.
NPTS	Number of observations in a data set.
RO	Values of the dependent variable (runoff).
PR	Values of the independent variable.

Output variable definitions

CASE A

A1	Sum of RO_{M-1}.
A2	Sum of RO_M.
A3	Sum of squares of RO_{M-1}.
A4	Sum of products of RO_{M-1} with RO_M.
A5	Sum of squares of RO_M.
ERROR	Observed minus predicted RO_M.
AA	Coefficient a.
BB	Coefficient c.

CASE B

A1	Sum of PR_M.
A2	Sum of RO_M.
A3	Sum of squares of PR_M.
A4	Sum of products of PR_M with RO_M.
A5	Sum of squares of RO_M.
ERROR	Observed minus predicted RO_M.
AA	Coefficient a.
BB	Coefficient b.

CASE C

A1	Sum of PR_M.
A2	Sum of RO_{M-1}.
A3	Sum of RO_M.
A4	Sum of squares of PR_M.
A5	Sum of products of PR_M with RO_{M-1}.
A6	Sum of products of PR_M with RO_M.
A7	Sum of squares of RO_{M-1}.
A8	Sum of products of RO_{M-1} with RO_M.
A9	Sum of squares of RO_M.
ERROR	Observed minus predicted RO_M.
AA	Coefficient a.
BB	Coefficient b.
CC	Coefficient c.

Note 1.—On second pass, values for cases A, B, and C are based on log-transformation of PR and RO.

Note 2.—For case A and case C, first output set is for February, and last output set is for January. For case B, first output set is for January, and last output set is for December.

```
C CALCULATION OF AUTO- AND CROSS-REGRESSION
C COEFFICIENTS BY MONTHS
      DIMENSION RO(1000),PR(1000),A1(12),A2(12),A3(12),A4(12),A5(12),A6(
     112),A7(12),A8(12),A9(12),ERROR(100)
      READ(5,1000) NSETS
      DO 1053 LSET=1,NSETS
      READ(5,1000) NPTS
```

```
 1000 FORMAT(I4)
      READ(5,1001) (RO(I),I=1,NPTS)
 1001 FORMAT(20F4.2)
      READ(5,1001) (PR(I),I=1,NPTS)
      ASSIGN 1020 TO KSPOT
      N=NPTS-1
C CASE A. BACK RUNOFF ONLY
 1022 DO 1002 II=1,12
      A1(II)=0.0
      A2(II)=0.0
      A3(II)=0.0
      A4(II)=0.0
 1002 A5(II)=0.0
      DO 1003 II=1,12
      DO 1004 I=II,N,12
      A1(II)=A1(II)+RO(I)
      A2(II)=A2(II)+RO(I+1)
      A3(II)=A3(II)+RO(I)*RO(I)
      A4(II)=A4(II)+RO(I+1)*RO(I)
 1004 A5(II)=A5(II)+RO(I+1)*RO(I+1)
 1003 CONTINUE
      WRITE(6,1007)
 1007 FORMAT(*0 VALUES FOR CASE A*)
      DO 1005 II=1,12
      WRITE(6,1006) II,A1(II),A2(II),A3(II),A4(II),A5(II)
 1006 FORMAT(*0*,I3,5E17.9)
 1005 CONTINUE
      DO 1028 II=1,12
      IF(II.EQ.12) GO TO 1026
      NUM=NPTS/12
      GO TO 1027
 1026 NUM=(NPTS/12)-1
 1027 X1=A1(II)/NUM
      X2=A3(II)/A1(II)
      X3=A2(II)/NUM
      X4=A4(II)/A1(II)
      BB=(X3-X4)/(X1-X2)
      AA=X3-(BB*X1)
      WRITE(6,1050)
 1050 FORMAT(*0*,/)
      IX=0
      DO 1030 IERR=II,N,12
      IX=IX+1
 1030 ERROR(IX)  =RO(IERR+1)-AA-(BB*RO(IERR))
      WRITE(6,1031) II,(IERR,ERROR(IERR),IERR=1,NJM)
 1031 FORMAT(*0ERRORS FOR*,I5/(8(I4,F10.7)))
      WRITE(6,1029) AA,BB
 1029 FORMAT(*0 AA IS*,F16.8,*   BB IS*,F16.8)
 1028 CONTINUE
C CASE B. CURRENT PRECIPITATION ONLY
      DO 1008 II=1,12
      A1(II)=0.0
      A2(II)=0.0
      A3(II)=0.0
      A4(II)=0.0
 1008 A5(II)=0.0
      DO 1009 II=1,12
      DO 1010 I=II,NPTS,12
```

```
      A1(II)=A1(II)+PR(I)
      A2(II)=A2(II)+RO(I)
      A3(II)=A3(II)+PR(I)*PR(I)
      A4(II)=A4(II)+PR(I)*RO(I)
 1010 A5(II)=A5(II)+RO(I)*RO(I)
 1009 CONTINUE
      WRITE(6,1011)
 1011 FORMAT(*0 VALUES FOR CASE B*)
      DO 1012 II=1,12
      WRITE(6,1006) II,A1(II),A2(II),A3(II),A4(II),A5(II)
 1012 CONTINUE
      DO 1032 II=1,12
      NUM=NPTS/12
      X1=A1(II)/NUM
      X2=A3(II)/A1(II)
      X3=A2(II)/NUM
      X4=A4(II)/A1(II)
      BB=(X3-X4)/(X1-X2)
      AA=X3-(BB*X1)
      WRITE(6,1050)
      IX=0
      DO 1033 IERR=II,NPTS,12
      IX=IX+1
 1033 ERROR(IX)  =RO(IERR)-AA-(BB*PR(IERR))
      WRITE(6,1031) II,(IERR,ERROR(IERR),IERR=1,NUM)
      WRITE(6,1029) AA,BB
 1032 CONTINUE
C CASE C. CURRENT PRECIPITATION AND BACK RUNOFF
      DO 1013 II=1,12
      A1(II)=0.0
      A2(II)=0.0
      A3(II)=0.0
      A4(II)=0.0
      A5(II)=0.0
      A6(II)=0.0
      A7(II)=0.0
      A8(II)=0.0
 1013 A9(II)=0.0
      DO 1014 II=1,12
      DO 1015 I=II,N,12
      A1(II)=A1(II)+PR(I+1)
      A2(II)=A2(II)+RO(I)
      A3(II)=A3(II)+RO(I+1)
      A4(II)=A4(II)+PR(I+1)*PR(I+1)
      A5(II)=A5(II)+PR(I+1)*RO(I)
      A6(II)=A6(II)+PR(I+1)*RO(I+1)
      A7(II)=A7(II)+RO(I)*RO(I)
      A8(II)=A8(II)+RO(I)*RO(I+1)
 1015 A9(II)=A9(II)+RO(I+1)*RO(I+1)
 1014 CONTINUE
      WRITE(6,1016)
 1016 FORMAT(*0 VALUES FOR CASE C*)
      DO 1025 II=1,12
      WRITE(6,1017) II,A1(II),A2(II),A3(II),A4(II),A5(II),A6(II),A7(II),
     1A8(II),A9(II)
 1017 FORMAT(*0*,I3,(5E17.9))
 1025 CONTINUE
      DO 1034 II=1,12
```

```
      IF(II.EQ.12) GO TO 1035
      NUM=NPTS/12
      GO TO 1036
1035  NUM=(NPTS/12)-1
1036  X1=A1(II)/NUM
      X2=A2(II)/NUM
      X3=A3(II)/NUM
      X4=A4(II)/A1(II)
      X5=A5(II)/A1(II)
      X6=A6(II)/A1(II)
      X7=A5(II)/A2(II)
      X8=A7(II)/A2(II)
      X9=A8(II)/A2(II)
      Y1=(X2-X5)/(X1-X4)
      Y2=(X3-X6)/(X1-X4)
      Y3=(X5-X8)/(X4-X7)
      Y4=(X6-X9)/(X4-X7)
      CC=(Y2-Y4)/(Y1-Y3)
      BB=Y2-(CC*Y1)
      AA=X3-(CC*X2)-(BB*X1)
      WRITE(6,1050)
      IX=0
      DO 1038 IERR=II,N,12
      IX=IX+1
1038  ERROR(IX)  =RO(IERR+1)-AA-(BB*PR(IERR+1))-(CC*RO(IERR))
      WRITE(6,1031) II,(IERR,ERROR(IERR),IERR=1,NUM)
      WRITE(6,1037) AA,BB,CC
1037  FORMAT(*0 AA IS*,F16.8,*   BB IS*,F16.8,*   CC IS*,F16.8)
1034  CONTINUE
      AUGMENT TO PREVENT LOG OF ZERO ERROR
      GO TO KSPOT,(1020,1021)
1020  ASSIGN 1021 TO KSPOT
      DO 1018 I=1,NPTS
      PR(I)=ALOG(PR(I)+0.10)
1018  RO(I)=ALOG(RO(I)+0.01)
      WRITE(6,1019)
1019  FORMAT(*0 PRECIP AND RUNOFF INCREASED,0.10 AND 0.01, RESPECTIVELY,
     1TO PREVENT LOG OF ZERO ERROR*/* FOLLOWING ARE BASED ON LOGS OF DAT
     2A*)
      WRITE(6,1051) (IKK,PR(IKK),IKK=1,NPTS)
1051  FORMAT(*0 LOGS OF RAINFALL PLUS 0.10*/(8(I5,F10.6)))
      WRITE(6,1052) (IKK,RO(IKK),IKK=1,NPTS)
1052  FORMAT(*0 LOGS OF RUNOFF PLUS 0.01*/(8(I5,F10.6)))
      GO TO 1022
1021  CONTINUE
1053  CONTINUE
      STOP
      END
```

Program III

This is a subprogram only, intended for use with some other program, or system of programs, to obtain optimum values of the coefficients. The program system used for this study was presented in DeCoursey and Snyder.[14]

To remove the restraint of negative values set to zero during fitting, the two statements labeled A can be deleted.

To incorporate the restraint of coefficient *b* not larger than 1.0, the short LIMIT subprogram can be used.

To fit log-transformed data, the four statements immediately following the program END can be inserted at point B.

Input variable definitions

OBSY — Values of runoff by months.

RAIN — Values of rainfall by months.

BRO — Back runoff. Runoff in the month immediately preceding the first OBSY of record utilized. It can be estimated.

Other input controls for the optimizing procedure used were read in by the system MAIN program.

Output variable definitions

Output is controlled by the system MAIN program or other subprograms. Normally, values of the coefficients, calculated values and errors for each observation, and sensitivity coefficients for each observation are output for selected rounds of iteration until solution converges to preset iteration limit.

[14]DeCoursey, D. G., and Snyder, W. M. 1969. Computer-oriented method of optimizing model parameters. J. Hydrol. 9(1): 34–56.

```
C THIS SUBPROGRAM IS USED WITH ITERATIVE NONLINEAR LEAST SQUARES
C TO EVALUATE SEASONALLY CONTINUOUS REGRESSION COEFFICIENTS.
      SUBROUTINE PARTAL(NOBS,NPAR,MRUN,KPRIN,ISTOP)
      COMMON/ONE/PAR(11),X(900,12),COE(11),OBSY(900),PREDY(900)/TWO/CORR(
     112,12),EVAL(11),FVEC(11,11)
      DIMENSION A(3,12), RAIN(500),RO(500),DR(11)
      IF(MRUN.NE.0) GO TO 1000
      READ(5,1001) (OBSY(I),I=1,NOBS)
      READ(5,1001)(RAIN(I),I=1,NOBS)
 1001 FORMAT(20F4.2)
      READ(5,1001) BRO                                               B
 1000 DO 1002 I=1,NPAR
      DR(I)=0.01*PAR(I)
      IF(ABS(DR(I)).LT.0.001) DR(I)=0.001
 1002 CONTINUE
      ASSIGN 2008 TO ITRAN
 1003 ASSIGN 1004 TO KSPCT
      P1=PAR(1)
      P2=PAR(2)
      P3=PAR(3)
      I=1
 1007 A(I,2)=(-0.0937*P3)+(0.6672*P1)+(0.2266*P2)
      A(I,3)=(-0.125*P3)+0.5625*(P1+P2)
      A(I,4)=(-0.0937*P3)+(0.2266*P1)+(0.6672*P2)
      A(I,5)=P2
      A(I,6)=(-0.0937*P1)+(0.6672*P2)+(0.2266*P3)
      A(I,7)=(-0.125*P1)+0.5625*(P2+P3)
      A(I,8)=(-0.0937*P1)+(0.2266*P2)+(0.6672*P3)
      A(I,9)=P3
      A(I,10)=(-0.0937*P2)+(0.6672*P3)+(0.2266*P1)
      A(I,11)=(-0.125*P2)+0.5625*(P3+P1)
      A(I,12)=(-0.0937*P2)+(0.2266*P3)+(0.6672*P1)
      A(I,1)=P1
      GO TO KSPCT,(1004,1005,1006)
 1004 ASSIGN 1005 TO KSPCT
      P1=PAR(4)
```

```
      P2=PAR(5)
      P3=PAR(6)
      I=2
      GO TO 1007
 1005 ASSIGN 1006 TO KSPOT
      P1=PAR(7)
      P2=PAR(8)
      P3=PAR(9)
      I=3
      GO TO 1007
 1006 RO(1)=A(1,1)+A(2,1)*RAIN(1)+A(3,1)*BRO
      IF(RO(1).LT.0.0) RO(1)=0.0                         A
      DO 1008 I=2,NOBS
      IN=MOD(I,12)
      IF(IN.EQ.0) IN=12
      RO(I)=A(1,IN)+A(2,IN)*RAIN(I)+A(3,IN)*OBSY(I-1)
      IF(RO(I).LT.0.0) RO(I)=0.0                            A
 1008 CONTINUE
      GO TO ITRAN,(2000,1015)
 2000 ASSIGN 1015 TO ITRAN
      DO 1016 J=1,NOBS
      X(J,10)=OBSY(J)-RO(J)
 1016 PREDY(J)=RO(J)
      ILP=1
 1021 SCOE=PAR(ILP)
      PAR(ILP)=PAR(ILP)+DP(ILP)
      GO TO 1003
 1015 PAR(ILP)=SCOE
      DO 1019 JK=1,NOBS
 1019 X(JK,ILP)=(RO(JK)-PREDY(JK))/DP(ILP)
      IF(ILP.EQ.9) GO TO 1020
      ILP=ILP+1
      GO TO 1021
 1020 RETURN
      END





      DO 3000 IS=1,NOBS
      OBSY(IS)=ALOG(OBSY(IS)+0.01)
 3000 RAIN(IS)=ALOG(RAIN(IS)+0.10)
      BRO=ALOG(BRO+0.01)





      SUBROUTINE LIMIT(NVAR,NRUN)
      COMMON/ONE/PAR(11),X( 900,12),COE(11),OBSY( 900),PREDY( 900)
      DO 1000 I=4,6
      IF(PAR(I).GT.1.0) GO TO 1001
      GO TO 1000
 1001 PAR(I)=1.0
      COE(I)=1.0-PAR(I)
 1000 CONTINUE
      RETURN
      END
```

Program IV

This is a subprogram only, intended for use with the same systems of programs as program III.

Input variable definitions

NYR	Number of years in the record.
NCLAS	Number of classes for each month.
WT	Value of exponential parameter for weighting errors.
OBSY	The number of observations in each class. Observations are read in as a continuous set and then organized into months by Do-Loop 3001.

Output variable definitions

Same as for program III.

```
C THIS SUBPROGRAM IS USED WITH ITERATIVE NONLINEAR LEAST SQUARES
C TO EVALUATE SEASONALLY CONTINUOUS DISTRIBUTION FUNCTIONS
      SUBROUTINE PARTAL(NOBS,NPAR,MRUN,KPRIN,ISTOP)
      COMMON/ONE/PAR(11),X(300,12),COE(11),OBSY(300),PREDY(300)/TWO/COR(
     112,12),EVAL(11),EVEC(11,11)
      DIMENSION DR(7),SOB(25,12),A(2,12),FUN(10),AVFUN(300)
      IF(MRUN.NE.0) GO TO 3000
      READ(5,6000) NYR,NCLAS,WT
 6000 FORMAT(2I4,F6.0)
      NOBS=12*NCLAS
      DO 5101 J=1,12
      IE=J*NCLAS
      IB=IE-NCLAS+1
 5101 READ(5,5100)(OBSY(I),I=IB,IE)
 5100 FORMAT(25F3.0)
      PIF=SQRT(2*3.1415927)
      DO 3001 I=1,12
      IB=(I-1)*NCLAS+1
      DO 3002 IS=1,NCLAS
 3002 SOB(IS,I)=OBSY(IB+IS-1)
 3001 CONTINUE
 3000 ASSIGN 4011 TO ITRAN
      ASSIGN 3018 TO JTRAN
      DO 4060 I=1,NPAR
      DR(I)=0.005*ABS(PAR(I))
      IF(DR(I).LT.0.005) DR(I)=0.005
 4060 CONTINUE
      DO 3010 I=1,300
      OBSY(I)=0.0
 3010 PREDY(I)=0.0
 3011 ASSIGN 3003 TO KSPOT
      P1=PAR(1)
      P2=PAR(2)
      P3=PAR(3)
      I=1
 3005 A(I,1)=P1
      A(I,2)=(-0.0937*P3)+(0.8672*P1)+(0.2266*P2)
      A(I,3)=(-0.125*P3)+0.5625*(P1+P2)
      A(I,4)=(-0.0937*P3)+(0.2266*P1)+(0.8672*P2)
      A(I,5)=P2
      A(I,6)=(-0.0937*P1)+(0.8672*P2)+(0.2266*P3)
      A(I,7)=(-0.125*P1)+0.5625*(P2+P3)
      A(I,8)=(-0.0937*P1)+(0.2266*P2)+(0.8672*P3)
      A(I,9)=P3
      A(I,10)=(-0.0937*P2)+(0.8672*P3)+(0.2266*P1)
      A(I,11)=(-0.125*P2)+0.5625*(P3+P1)
      A(I,12)=(-0.0937*P2)+(0.2266*P3)+(0.8672*P1)
      GO TO KSPOT,(3003,3004)
```

```
3003 ASSIGN 3004 TO KSPOT
     P1=PAR(4)
     P2=PAR(5)
     P3=PAR(6)
     I=2
     GO TO 3005
3004 NGTOE=0
     DO 3006 MON=1,12
     IADV=0
     CL=(-0.005)
     ICLASS=0
     IF(A(1,MON).LT.CL) GO TO 3007
     GO TO 5001
3008 IF(A(1,MON).LT.CL) GO TO 3007
     GO TO 3009
3007 IADV=IADV+1
     CL=CL-1.0
     GO TO 3008
5001 IF(A(1,MON).GT.(CL+1.0)) GO TO 5000
     GO TO 3009
5000 ICLASS=ICLASS+1
     IF(SOB(ICLASS,MON).GT.0) GO TO 3009
     IADV=IADV-1
     CL=CL+1.0
     GO TO 5001
3009 NTOT=NCLAS+IADV
     FUN0=0.0
     DO 3012 I=1,NTOT
     DO 3013 I4TH=1,10
     CL=CL+0.10
     IF(CL.LE.A(1,MON)) GO TO 3015
     GO TO 3016
3015 FUN(I4TH)=0.0
     GO TO 3013
3016 FUN(I4TH)=EXP(-0.5*((ALOG(CL-A(1,MON))/A(2,MON))-PAR(7))**2)
     FUN(I4TH)=FUN(I4TH)/(PIF*A(2,MON)*(CL-A(1,MON)))
3013 CONTINUE
     AVFUN(I) = 0.0
     DO 6101 II=1,9
6101 AVFUN(I) = AVFUN(I) + FUN(II)/10.0
     AVFUN(I) = AVFUN(I) + (FUN0 + FUN(10))/20.0
     AVFUN(I)=AVFUN(I)*NYR
3012 FUN0=FUN(10)
     GO TO ITRAN,(4011,4024)
4011 NGTSAV=NGTOE
     DO 3014 K=1,NTOT
     IX=NGTOE+K
3014 PREDY(IX)=AVFUN(K)
     IF(IADV.GE.0) GO TO 4100
     GO TO 4101
4100 IOB=NGTOE+1+IADV
     IOE=NGTOE+NTOT
     IC=1
     DO 4102 IO=IOB,IOE
     OBSY(IO)=SOB(IC,MON)
4102 IC=IC+1
     GO TO 4103
4101 IOB=NGTOE+1
```

```
      IOE=NGTOE+NTOT
      IC=IABS(IADV)+1
      DO 4104 IO=IOB,IOE
      OBSY(IO)=SOB(IC,MON)
4104  IC=IC+1
4103  CONTINUE
      DO 4016 L=1,NTOT
      IX=NGTOE+L
      IF(ISTOP.EQ.0) GO TO 1000
      GO TO 1001
1000  X(IX,8)=OBSY(IX)-PREDY(IX)
      GO TO 4016
1001  X(IX,8)=(OBSY(IX)-PREDY(IX))/(AVFUN(L)**WT)
4016  CONTINUE
      GO TO 3006
4024  DO 3017 M=1,NTOT
      IX=NGTOE+M
3017  X(IX,JCOE)=(AVFUN(M)-PREDY(IX))/DR(JCOE)
3006  NGTOE=NGTOE+NTOT
      GO TO JTRAN,(3018,3019)
3018  ASSIGN 3019 TO JTRAN
      ASSIGN 4024 TO ITRAN
      NOBS=NGTOE
      JCOE=1
3040  SCOE=PAR(JCOE)
      PAR(JCOE)=PAR(JCOE)+DR(JCOE)
      GO TO 3011
3019  PAR(JCOE)=SCOE
      IF(JCOE.EQ.NPAR) GO TO 3041
      JCOE=JCOE+1
      GO TO 3040
3041  RETURN
      END
```

Program V

Program V MAIN calls subroutine RANTAB. Subroutine RANTAB calls subroutine RNNG. Subroutine RNNG computes one random normal number. Subroutine RANTAB draws at random from a table of 100 random normal numbers with random replacement.[15] This method extends to almost inconceivable length the finite series of pseudorandom numbers before recycling begins. The integer numbers appearing in these subroutines are for a CDC 6400 computer.

Input variable definitions

STNS	The number of station simulations to be performed.
TITLE	Aphameric problem title.
NSETS	Number of sets (synthetic series) for each station simulation (for each set of coefficients and parameters).
NYRS	Number of years in each simulated series.
RN,N1,N2	Beginning random numbers.
BRO	Back runoff. Runoff in the month immediately preceding first month simulated.
CLASW	Class widths for the construction of histograms of monthly rainfall and runoff.
SDRES	Standard deviation of residuals, S_e.
BARERR	Average of residuals, $\overline{e}$.
COE	$a_1, a_5, a_9, b_1, b_5, b_9, c_1, c_5, c_9$.
PAR	$o_1, o_2, o_3, k_1, k_2, k_3, m$.

Output variable definitions

SUMP	Average of simulated monthly rainfall values.
SUMP2	Standard deviation of simulated monthly rainfall.
P	Values of simulated monthly rainfall.
SUMR	Average of simulated monthly runoff values.

[15]Grant, cited in footnote 11.

SUMR2 Standard deviation of simulated monthly runoff values.
RO Values of simulated monthly runoff.
NINCL Number of events in each class of the histograms of rainfall and runoff.
RN,N1,N2 Last values of random numbers generated. Can be input for subsequent runs.

```
C THIS PROGRAM SIMULATES MONTHLY RUNOFF IN INCHES USING SEASONALLY CYCLIC
C COEFFICIENTS OBTAIND BY METHOD OF CONTINUOUS PARABOLIC INTERPOLATION.
      DIMENSION X(12),A(12),B(12),C(12),D(12),E(12),COE(9),PAR(7),NINCL(
     1100,12),P(12,100),RO(12,100),TITLE(8),SUMP(12),SUMP2(12),SUMR(12),
     2SUMR2(12)
      INTEGER SET,YR, RN,STNS,STN
      READ(5,1042) STNS
 1042 FORMAT(I4)
      STN=1
 1044 READ(5,1037)(TITLE(I),I=1,8)
 1037 FORMAT( 8A10)
      WRITE(6,1038)(TITLE(I),I=1,8)
 1038 FORMAT(*1*,30X,8A10)
      READ(5,1000) NSETS,NYRS,RN,N1,N2
 1000 FORMAT(2I5,3I15)
      READ(5,1001) BRO,CLASW,SDRES,EARERR
 1001 FORMAT(10F8.0)
      READ(5,1001)  (COE(I),I=1,9)
      READ(5,1001)  (PAR(I),I=1,7)
      WRITE(6,1039) SDRES,BRO,RN,N1,N2
 1039 FORMAT(*0 STD. DEV. OF RANDOM ELEMENT*,F10.6/*  BACK RUNOFF IS*,F5
     1.2,* INCHES.*/*  BEGINNING RANDOM NOS. ARE*,3I20)
      WRITE(6,1040)(I,COE(I),I=1,9)
 1040 FORMAT(*0 THE COEFFICIENTS ARE...*/(9(I3,F10.6)))
      WRITE(6,1041)(I,PAR(I),I=1,7)
 1041 FORMAT(*0 THE DISTRIBUTION PARAMETERS FOR RAINFALL ARE...*/(7(I3,F
     110.6)))
C CALCULATE MONTHLY VALUES OF COEFFICIENT A.
      AA=COE(1)
      BB=COE(2)
      CC=COE(3)
      ASSIGN 1002 TO KSHIFT
 1005 X(1)=AA
      X(2)=(-0.0937*CC)+(0.8672*AA)+(0.2266*BB)
      X(3)=(-0.125*CC)+0.5625*(AA+BB)
      X(4)=(-0.0937*CC)+(0.2266*AA)+(0.8672*BB)
      X(5)=BB
      X(6)=(-0.0937*AA)+(0.8672*BB)+(0.2266*CC)
      X(7)=(-0.125*AA)+0.5625*(BB+CC)
      X(8)=(-0.0937*AA)+(0.2266*BB)+(0.8672*CC)
      X(9)=CC
      X(10)=(-0.0937*BB)+(0.8672*CC)+(0.2266*AA)
      X(11)=(-0.125*BB)+0.5625*(CC+AA)
      X(12)=(-0.0937*BB)+(0.2266*CC)+(0.8672*AA)
      GO TO KSHIFT,(1002,1004,1007,1009,1011)
 1002 DO 1003 I=1,12
 1003 A(I)=X(I)
      ASSIGN 1004 TO KSHIFT
```

```
C CALCULATE MONTHLY VALUES OF COEFFICIENT B.
      AA=COE(4)
      BB=COE(5)
      CC=COE(6)
      GO TO 1005
 1004 DO 1006 I=1,12
 1006 B(I)=X(I)
      ASSIGN 1007 TO KSHIFT
C CALCULATE MONTHLY VALUES OF COEFFICIENT C.
      AA=COE(7)
      BB=COE(8)
      CC=COE(9)
      GO TO 1005
 1007 DO 1008 I=1,12
 1008 C(I)=X(I)
C CALCULATE MONTHLY VALUES OF PARAMETER D.
      AA=PAR(1)
      BB=PAR(2)
      CC=PAR(3)
      ASSIGN 1009 TO KSHIFT
      GO TO 1005
 1009 DO 1010 I=1,12
 1010 D(I)=X(I)
      ASSIGN 1011 TO KSHIFT
C CALCULATE MONTHLY VALUES OF PARAMETER K.
      AA=PAR(4)
      BB=PAR(5)
      CC=PAR(6)
      GO TO 1005
 1011 DO 1012 I=1,12
 1012 E(I)=X(I)
      SET=1
C SET UP RANDOM NUMBER TABLE
C THEN CALCULATE PRECIP. AND RUNOFF BY MONTHS.
      CALL RANTAB(RN,N1,N2,DRAW,1)
 1034 DO 1013 YR=1,NYRS
      MON=1
 1014 CALL RANTAB(RN,N1,N2,DRAW,2)
      P(MON,YR)=(EXP(E(MON)*(DRAW+PAR(7)))+D(MON))*CLASW
      IF(P(MON,YR).LT.0.0) P(MON,YR)=0.0
      CALL RANTAB(RN,N1,N2,DRAW,2)
      ERR=DRAW*SDRES+BARERR
      RO(MON,YR)=A(MON)+B(MON)*P(MON,YR)+C(MON)*BRO+ERR
      BRO=RO(MON,YR)
      IF(RO(MON,YR).LT.0.0) RO(MON,YR)=0.0
      IF(MON.EQ.12) GO TO 1013
      MON=MON+1
      GO TO 1014
 1013 CONTINUE
C COMPUTE MEANS AND STANDARD DEVIATIONS
C OF SIMULATED VALUES BY CALENDAR MONTHS.
      DO 2501 MON=1,12
      SUMP(MON)=0.0
      SUMP2(MON)=0.0
      SUMR(MON)=0.0
      SUMR2(MON)=0.0
      DO 2500 YR=1,NYRS
      SUMP(MON)=SUMP(MON)+P(MON,YR)
```

```
      SUMP2(MON)=SUMP2(MON)+P(MON,YR)*P(MON,YR)
      SUMR(MON)=SUMR(MON)+RO(MON,YR)
 2500 SUMR2(MON)=SUMR2(MON)+RO(MON,YR)*RO(MON,YR)
      SUMP2(MON)=SQRT((SUMP2(MON)-(SUMP(MON)*SUMP(MON))/NYRS)/(NYRS-1))
      SUMR2(MON)=SQRT((SUMR2(MON)-(SUMR(MON)*SUMR(MON))/NYRS)/(NYRS-1))
      SUMP(MON)=SUMP(MON)/NYRS
 2501 SUMR(MON)=SUMR(MON)/NYRS
      DO 1015 MON=1,12
      WRITE(6,1016) MON,SUMP(MON),SUMP2(MON),(YR,P(MON,YR),YR=1,NYRS)
 1016 FORMAT(*0 RAINFALL FOR MONTH NO.*,I3/*  AVERAGE*,F7.3,*   STD. DEV
     1.*,F8.4/(12(I4,F6.2)))
 1015 WRITE(6,1017) MON,SUMR(MON),SUMR2(MON),(YR,RO(MON,YR),YR=1,NYRS)
 1017 FORMAT(*0 RUNOFF FOR MONTH NO.*,I3/*  AVERAGE*,F7.3,*  STD. DEV.*,
     1F8.4/(12(I4,F6.2)))
C CONSTRUCT HISTOGRAMS OF PRECIP. AND RUNOFF BY MONTHS.
      DO 1019 MON=1,12
      DO 1020 I=1,50
 1020 NINCL(I,MON)=0
      DO 1021 YR=1,NYRS
      CLAS=CLASW
      I=1
 1023 IF(P(MON,YR).LT.CLAS) GO TO 1022
      I=I+1
      CLAS=CLAS+CLASW
      GO TO 1023
 1022 NINCL(I,MON)=NINCL(I,MON)+1
 1021 CONTINUE
 1019 CONTINUE
      WRITE(6,1026) CLASW
 1026 FORMAT(*0 FOLLOWING HISTOGRAMS START AT 0.00 INCH AND HAVE A CLASS
     1WIDTH OF*,F5.2,* INCH.*)
      DO 1024 MON=1,12
 1024 WRITE(6,1025) MON,(I,NINCL(I,MON),I=1,50)
 1025 FORMAT(*0 HISTOGRAM OF RAINFALL FOR MONTH NO.*,I3/(10(I6,I4)))
      DO 1027 MON=1,12
      DO 1028 I=1,50
 1028 NINCL(I,MON)=0
      DO 1029 YR=1,NYRS
      CLAS=CLASW
      I=1
 1030 IF(RO(MON,YR).LT.CLAS) GO TO 1031
      I=I+1
      CLAS=CLAS+CLASW
      GO TO 1030
 1031 NINCL(I,MON)=NINCL(I,MON)+1
 1029 CONTINUE
 1027 CONTINUE
      DO 1032 MON=1,12
 1032 WRITE(6,1033) MON, (I,NINCL(I,MON),I=1,50)
 1033 FORMAT(*0 HISTOGRAM OF RUNOFF FOR MONTH NO.*,I3/(10(I6,I4)))
      IF(SET.EQ.NSETS) GO TO 1035
      SET=SET+1
      GO TO 1034
 1035 WRITE(6,1036) RN,N1,N2
 1036 FORMAT(*0 LAST RANDOM NUMBERS ARE*/3I17)
      IF(STN.EQ.STNS) GO TO 1043
      STN=STN+1
      GO TO 1044
 1043 STOP
      END
```

```
      SUBROUTINE RANTAB(RN,N1,N2,DRAW,IDENT)
      DIMENSION TAB(10,10)
      INTEGER RN
      IF(IDENT.NE.1) GO TO 1002
      K1=16777219
      K2=281474976710655
      II=INT(10.0*FLOAT(N1)/K2)+1
      N1=N1*K1
      JJ=INT(10.0*FLOAT(N2)/K2)+1
      N2=N2*K1
      DO 1000 I=1,10
      DO 1001 J=1,10
      CALL RNNG(RN,SUM)
 1001 TAB(I,J)=SUM
 1000 CONTINUE
      IC=1
 1002 DRAW=TAB(II,JJ)
      CALL RNNG(RN,SUM)
      TAB(II,JJ)=SUM
      IF(MOD(IC,2).EQ.0) GO TO 1003
      II=INT(10.0*FLOAT(N1)/K2)+1
      N1=N1*K1
      GO TO 1004
 1003 JJ=INT(10.0*FLOAT(N2)/K2)+1
      N2=N2*K1
 1004 IC=IC+1
      RETURN
      END
```

```
      SUBROUTINE RNNG(NP,SUM)
      SUM=0.0
      DO 10 I=1,12
      NP=NP*16777219
      RN=FLOAT(NP)/281474976710655
   10 SUM=SUM+RN
      SUM=SUM-6.0
      RETURN
      END
```


* 1976 - GPO - 1750-S /671 -583/24

U. S. DEPARTMENT OF AGRICULTURE
AGRICULTURAL RESEARCH SERVICE
SOUTHERN REGION
P. O. BOX 53326
NEW ORLEANS, LOUISIANA 70153

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U. S. DEPARTMENT OF
AGRICULTURE
AGR 101

Agricultural Research Service
UNITED STATES DEPARTMENT OF AGRICULTURE
in cooperation with
Georgia Institute of Technology
and
University of Georgia Agricultural Experiment Stations